

# EXHIBIT J



US005800988A

**United States Patent** [19]  
**Casterman et al.**

[11] **Patent Number:** **5,800,988**  
 [45] **Date of Patent:** **Sep. 1, 1998**

[54] **IMMUNOGLOBULINS DEVOID OF LIGHT CHAINS**

# OTHER PUBLICATIONS

[75] **Inventors:** Cecile Casterman; Raymond Hamers, both of Sint-Genesius-Rode, Belgium

Webster's II New Riverside University Dictionary. The Riverside Publishing Co./Houghton Mifflin Co., Boston MA, 1995.

[73] **Assignee:** Vrije Universiteit Brussel, Brussels, Belgium

Blier et al., "A Limited Number of B Cell Lineages Generates The Heterogeneity Of A Secondary Immune Response", *J. Immunology*, vol. 139, No. 12, pp. 3996-4006 (Dec. 15, 1987).

[21] **Appl. No.:** 467,282

Hamers-Casterman et al., "Nature", 363:446-48 (1993).

[22] **Filed:** Jun. 6, 1995

Roitt et al., "Immunology", Gower Medical Publishing, London, pp. 1.5 and 5.7 (1985).

## Related U.S. Application Data

[62] **Division of Ser. No.** 106,944, Aug. 17, 1993, abandoned.

Ward et al., "Nature", 341:544-546 (1989).

## Foreign Application Priority Data

Aug. 21, 1992 [EP] European Pat. Off. .... 92402326  
 May 21, 1993 [EP] European Pat. Off. .... 93401310

*Primary Examiner*—Frank C. Eisenschank

[51] **Int. Cl.<sup>6</sup>** ..... C12N 7/01; C12N 15/63; C07H 21/04

*Assistant Examiner*—Evelyn Rabin

[52] **U.S. Cl.** ..... 435/69.6; 435/235.1; 435/252.3; 435/320.1; 536/23.53

*Attorney, Agent, or Firm*—Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.

[58] **Field of Search** ..... 435/69.6, 235.1, 435/252.3, 320.1; 536/23.53

## [57] ABSTRACT

## [56] References Cited

There is provided an isolated immunoglobulin comprising two heavy polypeptide chains sufficient for the formation of a complete antigen binding site or several antigen binding sites, wherein the immunoglobulin is further devoid of light polypeptide chains.

## FOREIGN PATENT DOCUMENTS

WO 92/01787 2/1992 WIPO.

**6 Claims, 12 Drawing Sheets**

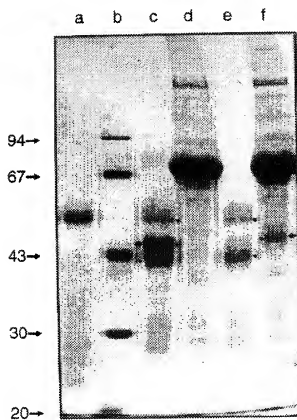


FIG. 1A

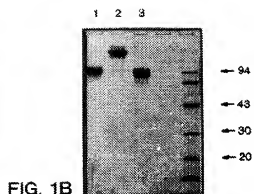


FIG. 1B

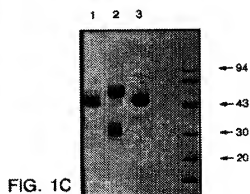


FIG. 1C

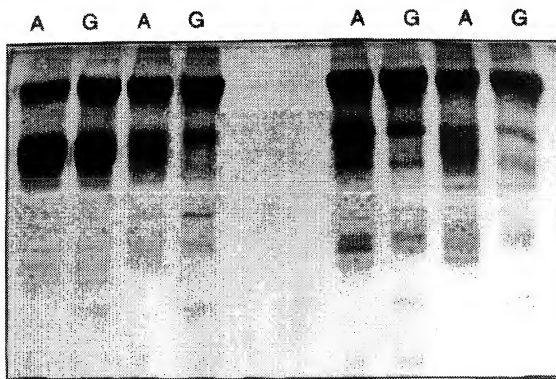


FIG. 2A

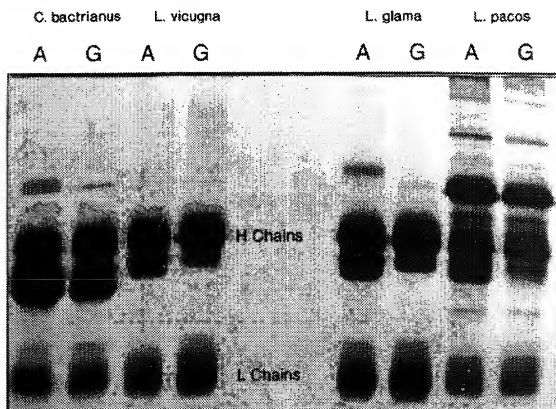
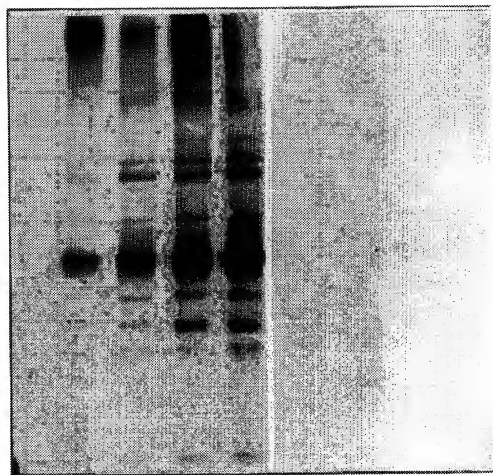


FIG. 2B



Prot. A	Ig1	Ig2	Ig3	Tot.Ser	Ig1	Ig2	Ig3	Tot.Ser	
Control	T. evansi infected				Healthy				
Counts/5ul	65	1258	1214	2700	2978	147	157	160	107

FIG. 3A

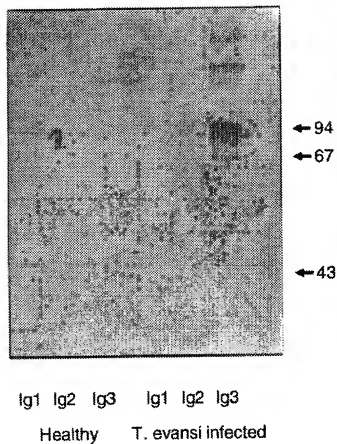


FIG. 3B



FIG. 3C

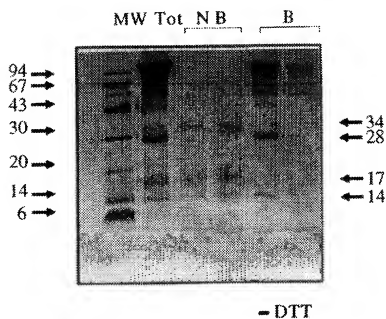


FIG. 4A

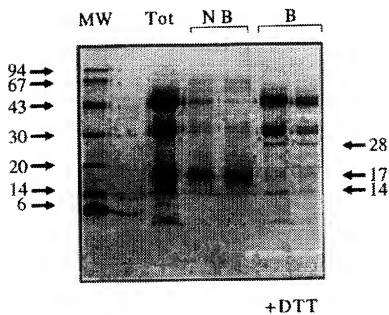
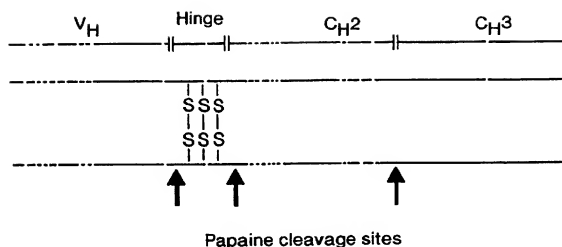


FIG. 4B



Papaine cleavage sites

FIG. 5



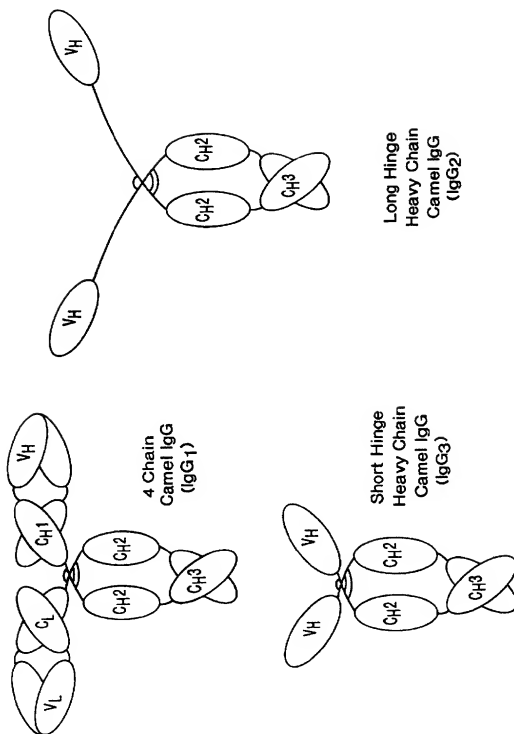


FIG. 6

```

DR01006 C-----TCGAG---TCTGGGGGAGG
DR27006 C-----TCGAG---TCTGGGGGAGG
DR03006 C-----AGGTGA-----AACTGCTCGAG---TCTGGAGGAGG
DR11006 C-----TCGAG---TCTGGGGGAGG
DR24006 C-----AGGTGA-----AACTGCTCGAG---TCTGGGGGAGG
DR16006 C-----TCGAG---TCTGGAGGAGG
DR19006 C-----TCGAG---TCTGGAGGAGG
DR07006 C-----TCGAG---TCTGGGGGAGG
DR16006 C-----TCGAG---TCTGGGGGAGG
DR20006 C-----TCGAG---TCAGGGGGAGG
DR25006 C-----TCGAG---TCTGGGGGAGG
DR20006 C-----TCGAG---TCTGGAGGAGG
DR21006 C-----TCGAG---TCTGGGGGAGG
DR09006 C-----AGGTGA-----AACTGCTCGAG---TCTGGGGGAGG
DR17006 C-----TCGAG---TCTGGGGGAGG
DR13006 C-----TCGAG---TCAGGGGGAGG
DR02006 CTCGAGTCAGGTGTCCGGTCTGATGTGACAGCTGGTGGCGTCTGGGGGAGG

DR01006 ATCGGTGCAGGCTGGAGGGTCTCTGAGACTCTC--GTGCG-CAGCCTCTG
DR27006 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTGCATCTTCTTCTA
DR03006 CTCGGTGCAGACTGGAGGATCTCTGAGACTCTCTGTGCACT--C-TCGTG
DR11006 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTAATGT--C-TCGTG
DR24006 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTAATGT--C-TCGTG
DR16006 CTCGGCGCAGGCTGGAGGATCTCTGAGACTCTCTGTGCAGC--CCACGG
DR19006 CTCGGTTTACGGCTGGAGGGTCCCTTAGACTCTCTGTGCAGC--C-TCGTG
DR07006 CTCGGTGCAGGGTGGAGGGTCTCTGAGACTCTCTGTGCAA---TCTCTG
DR16006 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTACAG--GCTCTG
DR20006 CTCGGTACAGGTTGGAGGGTCTCTGAGACTCTCTGTGTAG---CCTCTA
DR25006 CTCGGTACAACTGGAGGGTCTCTGAGACTCTCTTGC---AAATCTCTG
DR20006 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTG---TAGCCTCTG
DR21006 CTCGGTGCAGGTTGGAGGGTCTCTGAACTCTCTGTAAAAAT---CTCTG
DR09006 CTCGGTGCAGGCTGGAGGGTCTCTGACACTCTCTTGTG---TATACAC--
DR17006 CTCGGTCCAACCTGGAGGATCTCTGACACTCTCTGTACAGTT---TCTG
DR13006 CTCGGTGGAGGCTGGAGGGTCTCTGAGACTCTCTGTACAG---CCTCTG
DR02006 CTCGGTGCAGGCTGGAGGCTCTCTGAGACTCTCTGTACAG---CCTCTG

DR01006 GA--TACAGTAATT---GTCCCTCACTTG-GAGCTGGTATCGCCAGTTT
DR27006 AA--TATATGCCTT---GCACCTACGACAT-GACCTGGTACCGCCAGGCT
DR03006 GA--TTCTCCTTTA---GTACCACTTGTAT-GGCCTGGTTCGCCAGGCT
DR11006 GC--TCTCCAGTA--GTACTTATTGCCT-GGGCTGGTTCGCCAGGCT
DR24006 GC--TCTCCAGTA--GTACTTATTGCCT-GGGCTGGTTCGCCAGGCT
DR16006 GA--TTCCGC-TCA---ATGGTTACTACAT-CGCTGGTTCGGTCAGGCT
DR19006 AC--TACACCATCA---CTGATTATTGCAT-GGCCTGGTTCGCCAGGCT
DR07006 GA--TACACGTACG---GTAGCTTCTGTAT-GGCCTGGTTCGCCAGGCT
DR16006 GA--TTCCCCATA---GTACCTTCTGTCT-GGGGTGGTTCGCCAGGCT
DR20006 CT--CACACGCAGA---GTAGCACTGTAT-AGGCTGGTTCGCCAGGCT
DR25006 GA--TTGACTTTTG---ATGATTCTGACGT-GGGGTGGTACCGCCAGGCT
DR20006 GA--TTCAATTTTG---AACTTCTCGTAT-GGCCTGGTACCGCCAGCT
DR21006 GAGGTACCCAGATCGTGTCTCTAAATCTTTGGCTGGTTCGCCAGGCT
DR09006 -----CAACGATACTGGGACCA-----TGGGATGGTTCCGCCAGGCT
DR17006 ---GGGCCACCTACA---GTGACTACAATATTGGA-TGGATCCGCCAGGCT
DR13006 G-----ATACGTAT-CCT---CTATGGCTGGTTCGCCAGGCT
DR02006 GAGA----CAGTTTCAGTAGATT---TGCCATGTCTTGGTTCGCCAGGCT

```

FIG. 7A

DR01006 CCAGGAACGGAGCGCGAGTTCGTCTCCAGTATGGATCCGGATGGAAATAC  
DR27006 CCAGGCAAGGAGCGCGCAATTTGTCTCAAGTATAAATATTGATGGTAAGAC  
DR03006 TCAGGAAAGCAGCGTGAGGGGGTTCGACGCCATTAATAGTGCGGTGGTGA  
DR11006 CCAGGAGGGAGCGTGAGGGGGTTCACAGCGATTAA-----CACTGATGG  
DR24006 CCAGGGAAGGAGCGTGAGGGGGTTCACAGCGATTAA-----CACTGATGG  
DR16006 CCTGGGAAGGGCGTGAGGGGGTTCGCAACAATTAATGGTGGTCG-----  
DR19006 CCAGGGAAGGAGCGTGAAATTGGTTCGACGCGATTCAAGTTGTCCGTAGTGA  
DR07006 CCAGGCAAGGAACGTGAGGGGATCGCAACTATTCCTTAATGGTGGTACTAA  
DR16006 CCAGGGAAGGAGCGTGAGGGGGTTCGCGGGTATTAAATAGTGCAGAGGGTAA  
DR20006 CCAGGGAAGGAGCGCGAGGGGGTTCGCAAGTATATATTTTGGTGGTGGTGG  
DR25006 CCAGGGCATGAGTGCAAATTTGGTCTCAAGTATTCGAGTGATGGTACT-C  
DR20006 CCAGGAAATGTGTGTGAGTTGGTCTCAAGTATTACAGTGATGG-----  
DR21006 CCAGAGAAGGAGCGCGAGGGGATCGCAGTTCTTTTCGACTAAGGATGGTAA  
DR09006 CCAGGGAAGAGTGCAGAAAGGGTTCGCGCATATTACGCTGATGGTATGA  
DR17006 CCAGGGAAGGAGCGGTGAAGTAGTTCGACGCCCTAATACTGGTG-----  
DR13006 CCAGGGCAGGAGCGCGAGGGGGTTCGCGTTTGTCTCAAGCG-----  
DR20006 CCAGGGAAGGAGTGCAGAAATTGGTCTCAAGCATTCAAAGTATTGGAAGGAC

DR01006 CAAGTACA-----CATACTCCGTGAAGGGCCGCTTCAAC  
DR27006 AACATACG-----CAGACTCCGTGAAGGGCCGATTCAAC  
DR03006 GACATACTA-CAACACATATGTCGCCGAGTCCGTGAAGGGCCGATTCCGC  
DR11006 CAGTATCAT-ATACGCA-----GCCGACTCCGTGAAGGGCCGATTCAAC  
DR24006 CAGTGTCTAT-ATACGCA-----GCCGACTCCGTGAAGGGCCGATTCAAC  
DR16006 -----CGA-CGTACATACTACGCCGACTCCGTGAAGGGCCGATTCAAC  
DR19006 TACT-----CGC-C-TCACAGACTACGCCGACTCCGTGAAGGGCCGATTCAAC  
DR07006 -----CACATACTATGCCGACTCCGTGAAGGGCCGATTCAAC  
DR16006 -----TACTTACTATGCCGAGCGCGCTGAAGGGCCGATTCAAC  
DR20006 -----TACGAATTATCGCGACTCCGTGAAGGGCCGATTCAAC  
DR25006 CATATACAAGAGTGGAGACTATGCTGAGTCTGTGAGGGGCCGGGTACCA  
DR20006 CA-AAACATACTACGTCGACC-----GCA-----TGAAGGGCCGATTCAAC  
DR21006 GA-----CATTCTATGCCGACTCCGTGAAGGGCCGATTCAAC  
DR09006 -----CCTTCATTGATGAACCCGTGAAGGGCCGATTCAAC  
DR17006 ---CGACTAGTAAATTCACGTGACATTTGTGAAGGGCCGATTCAAC  
DR13006 ---CTGACAAT-AGTGCATTATAGGCGACTCCGTGAAGGGCCGATTCAAC  
DR20006 AACTGA-----GGCCGATTCCGTGAAGGGCCGATTCAAC

DR01006 ATGTCCCGAGGACAGCACCAGTACACAGTATTTCTGCAAAATGGACAATCT  
DR27006 ATCTCCCAAGACAGCGCCAAGAACACGGTGTATCTGACAGATGAACAGCCT  
DR03006 ATCTCCCAAGACAACGCCAAGACCACGGTATATCTGATATGAACAACCT  
DR11006 ATCTCCCAAGACACCGCCAAGGAACCGGTACATCTCGATGAACAACCT  
DR24006 ATCTCCCAAGACACCGCCAAGAAAACGGTATATCTCCAGATGAACAACCT  
DR16006 ATCTCCCGAGACAGCCCCAAGAATACGGTGTATCTGACAGTGAACAGCCT  
DR19006 ATCTCCCAAGGCAACACCAAGAACACAGTGAATCTGCAAAATGAACAGCCT  
DR07006 ATCTCCCAAGACAGCAGCTTGAAGACGATGTATCTGCAATGAACAACCT  
DR16006 ATCTCCCAAGGGAAATGCCAAGAATACGGTGTCTTCTGCAATGGATAACT  
DR20006 ATCTCCCAACTCAACGCCCAAGAACACAGTGTATCTGCAAAATGAACAACCT  
DR25006 ATCTCCAGAGACAACGCCAAGAACAATGATATACCTTCAATGAGCAACCT  
DR20006 ATTTCTAGAGAGAATGCCAAGAATACATTGTATCTTACAACAGAGCGCT  
DR21006 ATCTCTCTAGATAATGACAAGACCACTTTCTCCTTACAACCTGATCGACT  
DR09006 ATCTCCCGAGACAACGCCCAAGAAAACGTTGTCTTTCGGAATGAATAGTCT  
DR17006 ATTTCCCAAGACAACGCCAAGAATACGGTATATCTGCAAAATGAGCTTCT  
DR13006 ATCTCCCAAGACAACGCCAAGAACAACGCTGTATCTGCAAAATGCGCAACT  
DR20006 ATCTCCCGAGACAATTCCAGGAACACAGTGTATCTGCAAAATGAACAACCT

FIG. 7B

```

DR01006   GAAACCTGAGGACACGGCGATGTATTACTGTAAAC-A---GCCCTAC--
DR27006   GAAACCTGAGGACACGGCGATGTATTACTGTAAAT-A---GA---TTC-
DR03006   AACCCCTGAAGACACGGCTACGTATTACTGTGCGGCGG---TCCAGGCC
DR11006   GCAACCTGAGGATACGGCCACCTATTACTGCGCGGCCAA---GACTGACGG
DR24006   GCAACCTGAGGATACGGCCACCTATTACTGCGCGGCCAA---GACTGACGG
DR16006   GAAACCTGAGGACACGGCCATCTACTTCTGTGCGAGCAG---G-----CTC
DR19006   GACACCTGAGGACACGGCCATCTACAGTTGTGCGGCCAA---C-----CAG
DR07006   GAAACCTGAAGACACGGGCACCTATTACTGTGCTG-CA---GAACCTAAGT
DR16006   GAAACCTGAGGACACGGCCATCTATTACTGCGCGG-CG---GATAGTCCA
DR20006   GAAACCTGAGGACACGGCCATGTACTACTGTGCAATCA---GTAAATTTG
DR25006   GAAACCTGAGGACACGGCCATGTATTACTGCGCGGTAGATGGTTGGACCC
DR20006   AAAACCTGAGGACACGGCCATGTATTACTGTGCG---CC
DR21006   GAACCGGAGGACACTGCCGACTACTACTGCGCTGCAATCAATTAGC--
DR09006   GAGGCGCTGAGGACACGGCCGTGTATTACTGTGCGGCAGATTG---
DR17006   GAAACCTGAGGACACGGCCATCTATTACTGTGCGGCAG---CGGACCC
DR13006   GCAACCTGACGCACTGGCGTGTACTACTGTGCGGCC-----CAA
DR02006   GAAACCCGAGGACACGGCCGTGTATTACTGTGCGGCAGT-----

-----A-AC--CTGGGGTTATTGTGGGTA-
DR27006   -----GTAC--CCGTGCCATCTCCTTGATG-
DR03006   ACTTGGGACCT-----GGCG-CCATT-----CTTGATTTG
DR11006   AGATGGGGGCTTGTGATGCGAGATGGGCGACCTTAGC--GACAAGGAC-G
DR24006   AGATGGGGGCTTGTGATGCGAGATGGGCGACCTTAGC--GACAAGGAC-G
DR16006   GCGTTTTT-CTAGTCTGTGGGAGCACTTC-TAGAC---TCGAAAGTAG
DR19006   TAGTTTTTACTGTTACT-----ACG-----C---ACG---G
DR07006   GGTGGTAGTTGTGAAATTGC---CTTTGC-----TATTTGACTA-
DR16006   TGTTCATGCGGACTATGCG---CCGCTCCCCGATACGAGACAGTTTTGG
DR20006   AGTGGTATGGGTGCAATTT--AAGGACTACTTTTACT---C-----G
DR25006   GGAAGGAAG--GGGGAATCGGGTTAC-----CCTGGTCGGTCCAATGTGAA
DR20006   GGTTGAA-----TATC-----CTATTGCAGAC--ATGTGTT
DR21006   ---TGGTGGCTGGTATT-----TGGACCCGAATTACTGG-CTCTCTGTG
DR09006   ---GAAACTGGA-----CTTGTGGTGC--CCAGA-CTGG-----AG
DR17006   AAGTATATATTATAGTATC-----CTCCNNAT-----
DR13006   AAGAAGGATCGTA-----CTAGATGGGC-----CGAGCCT-----
DR02006   -----CTCCTCAA--TGGACCGAATTTG

--TGGGTANTGCCCTCTGGGGCCAGGGGACCCAGGTACCCGTCTCCTCACT
DR27006   --T-----CTGGGGCCAGGGGACCCAGGTACCCGTCTCCTCACT
DR03006   AAAAAGTATAAGTACTGGGGCCAGGGGACCCAGGTACCCGTCTCCTCACT
DR11006   TTTGCGTATAAATACTGGGGCCGGGGGACCCAGGTACCCGTCTCCTCACT
DR24006   TTTGCGTATAAATACTGGGGCCGGGGGACCCAGGTACCCGTCTCCTCACT
DR16006   CGA-CT-ATAAATAATTGGGGCCAGGGGATCCAAGTACCCGTCACTCACT
DR19006   CGC-CTTATAACGTCTGGGGTCAAGGGACCCAGGTACCCGTCTCCTCACT
DR07006   CTGGG-----GCCAGGGACCCAGGTACCCGTCTCCTCACT
DR16006   CTGGGATGATTTT-----GGCCAGGGGACCCAGGTACCCGTCTCCTCACT
DR20006   CTGGG-----GCCAGGGGACCCAGGTACCCGTCTCCTCACT
DR25006   GATGGTTATAAATAATTGGGGCCAGGGGACCCAGGTACCCGTCTCCTCACT
DR20006   CGAGAT-----ACG---GCGACCCGGGGACCCAGGTACCCGTCTCCTCACT
DR21006   GGTGCATATGCCATCTGGGGCCAGGGGACCCAGGTACCCGTCTCCTCACT
DR09006   GTACTTCCGACAG-TGGGGTCAGGGGGGACCCAGGTACCCGTCTCCTCACT
DR17006   --TGAGTATAAGTACTGGGGCCAGGGGACCCAGGTACCCGTCTCCTCACT
DR13006   CGAGAATGGAACAATACTGGGGCCAGGGGACCCAGGTACCCGTCTCCTCACT
DR02006   CCAACATGGG--TGCCGGGGCCAGGGGACCCAGGTACCCGTCTCCT--

```

FIG. 7C

DR01006 AG----TTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR27006 AG----TTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR03006 AGCTAGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR11006 AG----TTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR24006 AGCTAGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR16006 ----AGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR19006 ----AGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR07006 ----AGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR16006 ----AGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR20006 ----AGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR25006 ---TAGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR20006 ---TAGTTACCCGTACGACGAACCGGACTACGGTTCCTTAATAGAATTC  
DR21006 ---TAGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR09006 AGCTAGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR17006 -----  
DR13006 -----  
DR02006 -----TA

FIG. 7D

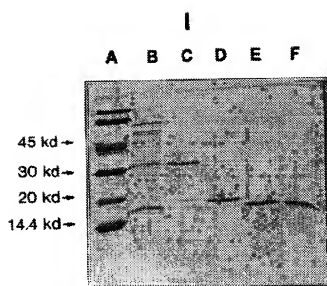


FIG. 8A

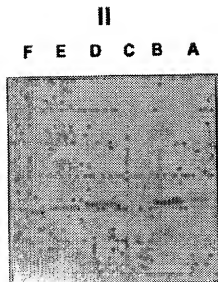


FIG. 8B

# IMMUNOGLOBULINS DEVOID OF LIGHT CHAINS

This is a division of application Ser. No. 08/106,944, filed Aug. 17, 1993 now abandoned.

The invention relates to new isolated immunoglobulins which are devoid of light polypeptide chains. These immunoglobulins do not consist in the degradation products of immunoglobulins composed of both heavy polypeptide and light polypeptide chains but to the contrary, the invention defines a new member of the family of the immunoglobulins, especially a new type of molecules capable of being involved in the immune recognition. Such immunoglobulins can be used for several purposes, especially for diagnosis or therapeutical purposes including protection against pathological agents or regulation of the expression or activity of proteins.

Up to now the structure proposed for immunoglobulins consists of a four-chain model referring to the presence of two identical light polypeptide chains (light chains) and two identical heavy polypeptide chains (heavy chains) linked together by disulfide bonds to form a Y- or T-shaped macromolecules. These chains are composed of a constant region and a variable region, the constant region being subdivided in several domains. The two heavy polypeptide chains are usually linked by disulfide bonds in a so-called "hinge region" situated between the first and second domains of the constant region.

Among the proteins forming the class of the immunoglobulins, most of them are antibodies and accordingly present an antigen binding site or several antigen binding sites.

According to the four-chain model, the antigen binding site of an antibody is located in the variable domains of each of the heavy and light chains, and requires the association of the heavy and the light chains variable domains.

For the definition of these four-chain model immunoglobulins, reference is made to Roitt, I et al (Immunology-second-Edition Gower Medical Publishing USA, 1989). Reference is especially made to the part concerning the definition of the four-chain immunoglobulins, their polypeptidic and genetic structures, the definition of their variable and constant regions and the obtention of the fragments produced by enzymatic degradation according to well known techniques.

The inventors have surprisingly established that different molecules can be isolated from animals which naturally produce them, which molecules have functional properties of immunoglobulins these functions being in some cases related to structural elements which are distinct from those involved in the function of four-chain immunoglobulins due for instance to the absence of light chains.

The invention relates to two-chain model immunoglobulins which neither correspond to fragments obtained for instance by the degradation in particular the enzymatic degradation of a natural four-chain model immunoglobulin, nor correspond to the expression in host cells, of DNA coding for the constant or the variable region of a natural four-chain model immunoglobulin or a part of these regions, nor correspond to antibodies produced in lymphocytes for example in mice, rats or human.

E. S. Ward et al (1) have described some experiments performed on variable domains of heavy polypeptide chains ( $V_H$ ) or/and light polypeptide chains ( $V_L$ ) to test the ability of these variable domains, to bind specific antigens. For this purpose, a library of  $V_H$  genes was prepared from the spleen genomic DNA of mice previously immunized with these specific antigens.

Ward et al have described in their publication that  $V_H$  domains are relatively sticky, presumably due to the exposed hydrophobic surface normally capped by the  $V_K$  or  $V_\lambda$  domains. They consequently envisage that it should be possible to design  $V_H$  domains having improved properties and further that  $V_H$  domains with binding activities could serve as the building blocks for making variable fragments (Fv fragments) or complete antibodies.

The invention does not start from the idea that the different fragments (light and heavy chains) and the different domains of these fragments of four-chain model immunoglobulin can be modified to define new or improved antigen binding sites or a four-chain model immunoglobulin.

The inventors have determined that immunoglobulins can have a different structure than the known four-chain model and that such different immunoglobulins offer new means for the preparation of diagnosis reagents, therapeutical agents or any other reagent for use in research or industrial purposes.

Thus the invention provides new immunoglobulins which are capable of showing functional properties of four-chain model immunoglobulins although their structure appears to be more appropriate in many circumstances for their use, their preparation and in some cases for their modification. Moreover these molecules can be considered as lead structures for the modification of other immunoglobulins. The advantages which are provided by these immunoglobulins comprise the possibility to prepare them with an increased facility.

The invention accordingly relates to immunoglobulins characterized in that they comprise two heavy polypeptide chains sufficient for the formation of a complete antigen binding site or several antigen binding sites, these immunoglobulins being further devoid of light polypeptide chains. In a particular embodiment of the invention, these immunoglobulins are further characterized by the fact that they are the product of the expression in a prokaryotic or in a eukaryotic host cell, of a DNA or of a cDNA having the sequence of an immunoglobulin devoid of light chains as obtainable from lymphocytes or other cells of Camels.

The immunoglobulins of the invention can be obtained for example from the sequences which are described in FIG. 7.

The immunoglobulins of the invention, which are devoid of light chains are such that the variable domains of their heavy chains have properties differing from those of the four-chain immunoglobulin  $V_H$ . The variable domain of a heavy-chain immunoglobulin of the invention has no normal interaction sites with the  $V_L$  or with the  $C_H1$  domain which do not exist in the heavy chain immunoglobulins. It is hence a novel fragment in many of its properties such as solubility and position of the binding site. For clarity reasons we will call it  $V_{HH}$  in this text to distinguish it from the classical  $V_H$  of four-chain immunoglobulins.

By "a complete antigen binding site" it is meant according to the invention, a site which will alone allow the recognition and complete binding of an antigen. This could be verified by any known method regarding the testing of the binding affinity.

These immunoglobulins which can be prepared by the technique of recombinant DNA, or isolated from animals, will be sometimes called "heavy-chain immunoglobulins" in the following pages. In a preferred embodiment of the invention, these immunoglobulins are in a pure form.

In a first embodiment, the immunoglobulins of the invention are obtainable in prokaryotic cells, especially in *E. coli* cells by a process comprising the steps of:

a) cloning in a Bluecript vector of a DNA or cDNA sequence coding for the  $V_{HH}$  domain of an immunoglobulin devoid of light chain obtainable for instance from lymphocytes of Camelids,

b) recovering the cloned fragment after amplification using a 5' primer containing an Xho site and a 3' primer containing the Spe site having the following sequence

TC TTA ACT AGT GAG GAG ACG GTG ACC TG, SEQ ID NO:31

c) cloning the recovered fragment in phase in the immuno PBS vector after digestion of the vector with Xho and Spe restriction enzymes,

d) transforming host cells, especially *E. coli* by transfection with the recombinant immuno PBS vector of step c,

e) recovering the expression product of the  $V_{HH}$  coding sequence, for instance by using antibodies raised against the dromadary  $V_{HH}$  domain.

In another embodiment the immunoglobulins are hetero-specific immunoglobulins obtainable by a process comprising the steps of:

obtaining a first DNA or cDNA sequence coding for a  $V_{HH}$  domain or part thereof having a determined specificity against a given antigen and comprised between Xho and Spe sites,

obtaining a second DNA or cDNA sequence coding for a  $V_{HH}$  domain or part thereof, having a determined specificity different from the specificity of the first DNA or cDNA sequence and comprised between the Spe and EcoRI sites,

digesting an immuno PBS vector with EcoRI and XhoI restriction enzymes,

ligating the obtained DNA or cDNA sequences coding for  $V_{HH}$  domains, so that the DNA or cDNA sequences are serially cloned in the vector,

transforming a host cell, especially *E. coli* cell by transfection, and recovering the obtained immunoglobulins.

In another embodiment, the immunoglobulins are obtainable by a process comprising the steps of:

obtaining a DNA or cDNA sequence coding for a  $V_{HH}$  domain or part thereof, having a determined specific antigen binding site,

amplifying the obtained DNA or cDNA, using a 5' primer containing an initiation codon and a HindIII site, and a 3' primer containing a termination codon having a XhoI site,

recombining the amplified DNA or cDNA into the HindIII (position 2650) and XhoI (position 4067) sites of a plasmid pMM984,

transfecting permissive cells especially NB-E cells with the recombinant plasmid,

recovering the obtained products.

Successful expression can be verified with antibodies directed against a region of a  $V_{HH}$  domain, especially by an ELISA assay.

According to another particular embodiment of this process, the immunoglobulins are cloned in a parvovirus.

In another example these immunoglobulins are obtainable by a process comprising the further cloning of a second DNA or cDNA sequence having another determined antigen binding site, in the pMM984 plasmid.

Such an Immunoglobulin can be further characterized in that it is obtainable by a process wherein the vector is Yep 52 and the transformed recombinant cell is a yeast especially *S. cerevisiae*.

A particular Immunoglobulin is characterized in that it has a catalytic activity, especially in that it is directed against an antigen mimicking an activated state of a given substrate. These catalytic antibodies can be modified at the level of their binding site, by random or directed mutagenesis in order to increase or modify their catalytic function. Reference may be made to the publication of Lerner et al (TIBS November 1987, 427-430) for the general technique for the preparation of such catalytic immunoglobulins.

According to a preferred embodiment, the immunoglobulins of the invention are characterized in that their variable regions contain in position 45, an amino-acid which is different from leucine, proline or glutamine residue.

Moreover the heavy-chain immunoglobulins are not products characteristic of lymphocytes of animals nor from lymphocytes of a human patient suffering from lymphopathies. Such immunoglobulins produced in lymphopathies are monoclonal in origin and result from pathogenic mutations at the genomic level. They have apparently no antigen binding site.

The two heavy polypeptide chains of these immunoglobulins can be linked by a hinge region according to the definition of Roitt et al.

In a particular embodiment of the invention, immunoglobulins corresponding to the above-defined molecules are capable of acting as antibodies.

The antigen binding site(s) of the immunoglobulins of the invention are located in the variable region of the heavy chain.

In a particular group of these immunoglobulins each heavy polypeptide chain contains one antigen binding site on its variable region, and these sites correspond to the same amino-acid sequence.

In a further embodiment of the invention the immunoglobulins are characterized in that their heavy polypeptide chains contain a variable region ( $V_{HH}$ ) and a constant region ( $C_H$ ) according to the definition of Roitt et al, but are devoid of the first domain of their constant region. This first domain of the constant region is called  $C_{H1}$ .

These immunoglobulins having no  $C_{H1}$  domain are such that the variable region of their chains is directly linked to the hinge region at the C-terminal part of the variable region.

The immunoglobulins of the type described hereabove can comprise type G immunoglobulins and especially immunoglobulins which are defined as immunoglobulins of class 2 (IgG2) or immunoglobulins of class 3 (IgG3).

The absence of the light chain and of the first constant domain lead to a modification of the nomenclature of the immunoglobulin fragments obtained by enzymatic digestion, according to Roitt et al.

The terms Fc and pFc on the one hand, Fc' and pFc' on the other hand corresponding respectively to the papain and pepsin digestion fragments are maintained.

The terms Fab F(ab)<sub>2</sub> F(ab')<sub>2</sub> Fabc, Fd and Fv are no longer applicable in their original sense as these fragments have either a light chain, the variable part of the light chain or the  $C_{H1}$  domain.

The fragments obtained by papain digestion and composed of the  $V_{HH}$  domain and the hinge region will be called FV<sub>HH</sub>h or F(V<sub>HH</sub>h)<sub>2</sub> depending upon whether or not they remain linked by the disulphide bonds.

In another embodiment of the invention, immunoglobulins replying to the hereabove given definitions can be originating from animals especially from animals of the camelid family. The inventors have found out that the heavy-chain immunoglobulins which are present in camelids are not associated with a pathological situation which



would induce the production of abnormal antibodies with respect to the four-chain immunoglobulins. On the basis of a comparative study of old world camelids (*Camelus bactrianus* and *Camelus dromaderius*) and new world camelids (for example *Lama pacos*, *Lama glama*, and *Lama vicugna*) the inventors have shown that the immunoglobulins of the invention, which are devoid of light polypeptide chains are found in all species. Nevertheless differences may be apparent in molecular weight of these immunoglobulins depending on the animals. Especially the molecular weight of a heavy chain contained in these immunoglobulins can be from approximately 43 kd to approximately 47 kd, in particular 45 kd.

Advantageously the heavy-chain immunoglobulins of the invention are secreted in blood of camelids.

Immunoglobulins according to this particular embodiment of the invention are obtainable by purification from serum of camelids and a process for the purification is described in details in the examples. In the case where the immunoglobulins are obtained from Camelids, the invention relates to immunoglobulins which are not in their natural biological environment.

According to the invention immunoglobulin IgG2, as obtainable by purification from the serum of camelids can be characterized in that:

it is not adsorbed by chromatography on Protein G Sepharose column,

it is adsorbed by chromatography on Protein A Sepharose column,

it has a molecular weight of around 100 kd after elution with a pH 4.5 buffer (0.15M NaCl, 0.58% acetic acid adjusted to pH 4.5 by NaOH),

has a molecular weight of around 100 kd after elution with a pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid),

is adsorbed by chromatography on a Protein G Sepharose column and eluted with pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid).

consists of heavy  $\gamma 3$  polypeptide chains of a molecular weight of around 45 Kd in particular between 43 and 47 kd after reduction.

The immunoglobulins of the invention which are devoid of light chains, nevertheless comprise on their heavy chains a constant region and a variable region. The constant region comprises different domains.

The variable region of immunoglobulins of the invention comprises frameworks (FW) and complementarity determining regions (CDR), especially 4 frameworks and 3 complementarity regions. It is distinguished from the four-chain immunoglobulins especially by the fact that this variable region can itself contain an antigen binding site or several, without contribution of the variable region of a light chain which is absent.

The amino-acid sequences of frameworks 1 and 4 comprise among others respectively amino-acid sequences which can be selected from the following:

for the framework 1 domain

G	G	S	V	Q	T	G	G	S	L	R	L	S	C	E	I	S	G	L	T	F	D	SEQ ID NO:1
G	G	S	V	Q	T	G	G	S	L	R	L	S	C	A	V	S	G	F	S	F	S	SEQ ID NO:2
G	G	S	E	Q	G	G	G	S	L	R	L	S	C	A	I	S	G	Y	T	Y	G	SEQ ID NO:3
G	G	S	V	Q	P	G	G	S	L	T	L	S	C	T	V	S	G	A	T	Y	S	SEQ ID NO:4
G	G	S	V	Q	A	G	G	S	L	R	L	S	C	V	A	G	F	G	T	S	S	SEQ ID NO:5
G	G	S	V	Q	A	G	G	S	L	R	L	S	C	V	S	F	S	P	S	S	SEQ ID NO:6	
G	G	S	V	Q	A	G	G	S	L	R	L	S	C	V	S	F	S	P	S	S	SEQ ID NO:7	

it consists of heavy  $\gamma 2$  polypeptide chains of a molecular weight of around 46 kd preferably 45 after reduction.

According to a further embodiment of the invention another group of immunoglobulins corresponding to IgG3, as obtainable by purification from the serum of Camelids is characterized in that the immunoglobulin:

is adsorbed by chromatography on a Protein A Sepharose column,

for the framework 4 domain

W	G	Q	G	T	Q	V	T	V	S	S	SEQ ID NO:8
W	G	Q	G	T	L	V	T	V	S	S	SEQ ID NO:9
W	G	Q	G	A	Q	V	T	V	S	S	SEQ ID NO:10
W	G	Q	G	T	Q	V	T	A	S	S	SEQ ID NO:11
R	G	Q	G	T	Q	V	T	V	S	L	SEQ ID NO:12

for the CDR3 domain

A	L	Q	P	G	Y	C	G	Y	G	X	—	—	—	—	—	—	—	—	—	C	L	SEQ ID NO:62		
V	S	L	M	D	R	I	S	Q	H	—	—	—	—	—	—	—	—	—	—	G	C	SEQ ID NO:63		
V	P	A	H	L	G	P	G	A	I	L	D	L	K	K	Y	—	—	—	—	K	Y	SEQ ID NO:64		
F	C	Y	S	T	A	G	D	G	G	S	G	E	—	—	—	—	—	—	—	M	Y	SEQ ID NO:65		
E	L	S	G	S	C	E	L	P	L	L	F	—	—	—	—	—	—	—	—	D	Y	SEQ ID NO:66		
R	L	W	K	Y	W	T	C	G	A	Q	T	G	Y	F	—	—	—	—	—	G	Q	SEQ ID NO:67		
R	L	T	E	M	G	A	C	D	A	R	W	A	T	L	A	T	R	T	F	A	Y	N	Y	SEQ ID NO:68
Q	K	K	D	R	T	R	W	A	E	P	H	W	—	—	—	—	—	—	—	N	N	SEQ ID NO:69		
G	S	R	P	S	S	P	V	G	S	T	S	R	L	E	S	—	S	D	Y	—	N	Y	SEQ ID NO:70	
A	D	P	S	I	Y	S	I	L	X	I	E	Y	—	—	—	—	—	—	—	K	Y	SEQ ID NO:71		
D	S	P	C	Y	M	P	T	M	P	A	P	T	Y	F	L	S	V	G	W	—	D	D	SEQ ID NO:72	
T	S	P	Y	W	Y	C	T	T	A	P	Y	—	—	—	—	—	—	—	—	N	V	SEQ ID NO:73		
T	E	I	E	W	Y	G	C	N	L	R	T	T	Y	F	L	S	V	G	A	Y	—	T	R	SEQ ID NO:74
N	Q	L	A	G	Q	W	Y	L	D	P	N	Y	—	—	—	—	—	—	—	A	I	SEQ ID NO:75		
R	L	T	E	M	G	A	C	D	A	R	W	A	T	L	A	T	R	T	F	A	Y	N	Y	SEQ ID NO:76
D	G	W	T	R	K	E	G	G	I	G	L	P	W	S	V	Q	C	E	D	G	Y	N	Y	SEQ ID NO:77





the genetic code. They can be synthesized or isolated from cells producing immunoglobulins of the invention.

A procedure for the obtention of such DNA sequences is described in the examples.

The invention also contemplates RNA, especially mRNA sequences corresponding to these DNA sequences, and also corresponding cDNA sequences.

The nucleotide sequences of the invention can further be used for the preparation of primers appropriate for the detection in cells or screening of DNA or cDNA libraries to isolate nucleotide sequences coding for immunoglobulins of the invention.

Such nucleotide sequences can be used for the preparation of recombinant vectors and the expression of these sequences contained in the vectors by host cells especially prokaryotic cells like bacteria or also eukaryotic cells and for example CHO cells, insect cells, simian cells like Vero cells, or any other mammalian cells. Especially the fact that the immunoglobulins of the invention are devoid of light chains permits to secrete them in eukaryotic cells since there is no need to have recourse to the step consisting in the formation of the BIP protein which is required in the four-chain immunoglobulins.

The inadequacies of the known methods for producing monoclonal antibodies or immunoglobulins by recombinant DNA technology comes from the necessity in the vast majority of cases to clone simultaneously the  $V_H$  and  $V_L$  domains corresponding to the specific binding site of 4 chain immunoglobulins. The animals and especially camels which produce heavy-chain immunoglobulins according to the invention, and possibly other vertebrate species are capable of producing heavy-chain immunoglobulins of which the binding site is located exclusively in the  $V_H$  domain. Unlike the few heavy-chain immunoglobulins produced in other species by chain separation or by direct cloning, the camelid heavy-chain immunoglobulins have undergone extensive maturation in vivo. Moreover their V region has naturally evolved to function in absence of the  $V_L$ . They are therefore ideal for producing monoclonal antibodies by recombinant DNA technology. As the obtention of specific antigen binding clones does not depend on a stochastic process necessitating a very large number of recombinant cells, this allows also a much more extensive examination of the repertoire.

This can be done at the level of the non rearranged  $V_H$  repertoire using DNA derived from an arbitrarily chosen tissue or cell type or at the level of the rearranged  $V_H$  repertoire, using DNA obtained from B lymphocytes. More interesting however is to transcribe the mRNA from antibody producing cells and to clone the cDNA with or without prior amplification into an adequate vector. This will result in the obtention of antibodies which have already undergone affinity maturation. The examination of a large repertoire should prove to be particularly useful in the search for antibodies with catalytic activities.

The invention thus provides libraries which can be generated in a way which includes part of the hinge sequence, the identification is simple as the hinge is directly attached to the  $V_H$  domain.

These libraries can be obtained by cloning cDNA from lymphoid cells with or without prior PCR amplification. The PCR primers are located in the promoter, leader or framework sequences of the  $V_H$  for the 5' primer and in the hinge,  $C_{H2}$ ,  $C_{H3}$ , 3' untranslated region or polyA tail for the 3' primer. A size selection of amplified material allows the construction of a library limited to heavy chain immunoglobulins.

In a particular example, the following 3' primer in which a KpnI site has been constructed and which corresponds to amino-acids 313 to 319 (CGC CAT CAA GGT AAC AGT TGA) SEQ ID NO:47 is used in conjunction with mouse  $V_H$  primers described by Sestry et al and containing a Xho site

AG	GTC	CAG	CTG	CTC	GAG	TCT	GG	SEQ ID NO:48
AG	CTC	CAG	CTG	CTC	GAG	TCT	GG	SEQ ID NO:49
AG	GTC	CAG	CTT	CTC	GAG	TCT	GG	SEQ ID NO:50

XhoI site

These primers yield a library of camelid heavy chain immunoglobulins comprising the  $V_H$  region (related to mouse or human subgroup III), the hinge and a section of  $CH_1$ .

In another example, the cDNA is polyadenylated at its 5' end and the mouse specific  $V_H$  primers are replaced by a poly T primer with an inbuilt XhoI site, at the level of nucleotide 12.

CTCGAGT<sub>12</sub>

The same 3' primer with a KpnI site is used.

This method generates a library containing all subgroups of immunoglobulins.

Part of the interest in cloning a region encompassing the hinge-CH<sub>2</sub> link is that in both  $\gamma 2$  and  $\gamma 3$ , a Sac site is present immediately after the hinge. This site allows the grafting of the sequence coding for the  $V_H$  and the hinge onto the Fc region of other immunoglobulins, in particular the human IgG<sub>1</sub> and IgG<sub>3</sub> which have the same amino acid sequence at this site (Glu<sub>346</sub> Leu<sub>347</sub>). As an example, the invention contemplates a cDNA library composed of nucleotide sequences coding for a heavy-chain immunoglobulin, such as obtained by performing the following steps:

- a) treating a sample containing lymphoid cells, especially peripheral, lymphocytes, spleen cells, lymph nodes or another lymphoid tissue from a healthy animal, especially selected among the Camelids, in order to separate the lymphoid cells,
- b) separating polyadenylated RNA from the other nucleic acids and components of the cells,
- c) reacting the obtained RNA with a reverse transcriptase in order to obtain the corresponding cDNA,
- d) contacting the cDNA of step c) with 5' primers corresponding to mouse  $V_H$  domain of four-chain immunoglobulins, which primer contains a determined restriction site, for example an XhoI site and with 3' primers corresponding to the N-terminal part of a  $C_{H2}$  domain containing a KpnI site,
- e) amplifying the DNA,
- f) cloning the amplified sequence in a vector, especially in a bluescript vector,
- g) recovering the clones hybridizing with a probe corresponding to the sequence coding for a constant domain from an isolated heavy-chain immunoglobulin.

This cloning gives rise to clones containing DNA sequences including the sequence coding for the hinge. It thus permits the characterization of the subclass of the immunoglobulin and the SacI site useful for grafting the FV<sub>H</sub>h to the Fc region.

The recovery of the sequences coding for the heavy-chain immunoglobulins can also be achieved by the selection of clones containing DNA sequences having a size compatible with the lack of the  $C_{H2}$  domain.

It is possible according to another embodiment of the invention, to add the following steps between steps c) and d) of the above process:

in the presence of a DNA polymerase and of deoxyribonucleotide triphosphates, contacting said cDNA with oligonucleotide degenerated primers, which sequences are capable of coding for the hinge region and N-terminal  $V_{HH}$  domain of an immunoglobulin, the primers being capable of hybridizing with the cDNA and capable of initiating the extension of a DNA sequence complementary to the cDNA used as template,

recovering the amplified DNA.

The clones can be expressed in several types of expression vectors. As an example using a commercially available vector Immuno PBS (Huse et al: Science (1989) 246, 1275), clones produced in Bluescript® according to the above described procedure, are recovered by PCR using the same XhoI containing 5' primer and a new 3' primer, corresponding to residues 113-103 in the framework of the immunoglobulins, in which an Spe site has been constructed: TC TTA ACT AGT GAG GAG ACG GTG ACC TG SEQ ID NO:51. This procedure allows the cloning of the  $V_{HH}$  in the Xho/Spe site of the Immuno PBS vector. However, the 3' end of the gene is not in phase with the identification "tag" and the stop codon of the vector. To achieve this, the construct is cut with Spe and the 4 base overhangs are filled in, using the Klenow fragment after which the vector is religated. A further refinement consists in replacing the marker ("tag") with a poly histidine so that metal purification of the cloned  $V_{HH}$  can be performed. To achieve this a Spe/EcoRI double stranded oligonucleotide coding for 6 histidines and a termination codon is first constructed by synthesis of both strands followed by heating and annealing:

CTA GTG CAC CAC CAT CAC CAT CAC TAA\* TAG\* SEQ ID NO:52  
AC GTG GTG GTA GTG GTA GTG ATT ATC TTA SEQ ID NO:53

The vector containing the insert is then digested with SpeI and EcoRI to remove the resident "tag" sequence which can be replaced by the poly-His/termination sequence. The produced  $V_{HH}$  can equally be detected by using antibodies raised against the dromedary  $V_{HH}$  regions. Under laboratory conditions,  $V_{HH}$  regions are produced in the Immuno PBS vector in mg amounts per liter.

The invention also relates to a DNA library composed of nucleotide sequences coding for a heavy-chain immunoglobulin, such as obtained from cells with rearranged immunoglobulin genes.

In a preferred embodiment of the invention, the library is prepared from cells from an animal previously immunized against a determined antigen. This allows the selection of antibodies having a preselected specificity for the antigen used for immunization.

In another embodiment of the invention, the amplification of the cDNA is not performed prior to the cloning of the cDNA.

The heavy-chain of the four-chain immunoglobulins remains sequestered in the cell by a chaperon protein (BiP) until it has combined with a light chain. The binding site for the chaperon protein is the  $C_{H1}$  domain. As this domain is absent from the heavy chain immunoglobulins, their secretion is independent of the presence of the BiP protein or of the light chain. Moreover the inventors have shown that the obtained immunoglobulins are not sticky and accordingly will not abnormally aggregate.

The invention also relates to a process for the preparation of a monoclonal antibody directed against a determined antigen, the antigen binding site of the antibody consisting of heavy polypeptide chains and which antibody is further devoid of light polypeptide chains, which process comprises:

immortalizing lymphocytes, obtained for example from the peripheral blood of Camelids previously immunized with a determined antigen, with an immortal cell and preferably with myeloma cells, in order to form a hybridoma,

culturing the immortalized cells (hybridoma) formed and recovering the cells producing the antibodies having the desired specificity.

The preparation of antibodies can also be performed without a previous immunization of Camelids.

According to another process for the preparation of antibodies, the recourse to the technique of the hybridoma cell is not required.

According to such process, antibodies are prepared in vitro and they can be obtained by a process comprising the steps of:

cloning into vectors, especially into phages and more particularly filamentous bacteriophages, DNA or cDNA sequences obtained from lymphocytes especially PBLs of Camelids previously immunized with determined antigens,

transforming prokaryotic cells with the above vectors in conditions allowing the production of the antibodies, selecting the antibodies for their heavy-chain structure and further by subjecting them to antigen-affinity selection,

recovering the antibodies having the desired specificity,

In another embodiment of the invention the cloning is performed in vectors, especially into plasmids coding for

bacterial membrane proteins. Prokaryotic cells are then transformed with the above vectors in conditions allowing the expression of antibodies in their membrane.

The positive cells are further selected by antigen affinity selection.

The heavy chain antibodies which do not contain the  $C_{H1}$  domain present a distinct advantage in this respect. Indeed, the  $C_{H1}$  domain binds to BiP type chaperone proteins present within eukaryotic vectors and the heavy chains are not transported out of the endocytosolic reticulum unless light chains are present. This means that in eukaryotic cells, efficient cloning of 4-chain immunoglobulins in non mammalian cells such as yeast cells can depend on the properties of the resident BiP type chaperone and can hence be very difficult to achieve. In this respect the heavy chain antibodies of the invention which lack the  $C_{H1}$  domain present a distinctive advantage.

In a preferred embodiment of the invention the cloning can be performed in yeast either for the production of antibodies or for the modification of the metabolism of the yeast. As example, Yep 52 vector can be used. This vector has the origin of replication (ORI) 2 $\mu$  of the yeast together with a selection marker Leu 2.

The cloned gene is under the control of gal promoter and accordingly is inducible by galactose. Moreover, the expression can be repressed by glucose which allows the obtention of very high concentration of cells before the induction.

The cloning between BamHI and SalI sites using the same strategy of production of genes by PCR as the one described above, allows the cloning of camelid immunoglobulin genes in *E. coli*. As example of metabolic modulation which can be obtained by antibodies and proposed for the yeast, one can site the cloning of antibodies directed against cyclins, that is proteins involved in the regulation of the cellular cycle of the yeast (TIBS 16 430 J. D. Mc Kinney, N. Heintz 1991). Another example is the introduction by genetic engineering of an antibody directed against CD<sub>28</sub>, which antibody would be inducible (for instance by gall), within the genome of the yeast. The CD<sub>28</sub> is involved at the level of the initiation of cell division, and therefore the expression of antibodies against this molecule would allow an efficient control of multiplication of the cells and the optimization of methods for the production in bioreactors or by means of immobilized cells.

In yet another embodiment of the invention, the cloning vector is a plasmid or a eukaryotic virus vector and the cells to be transformed are eukaryotic cells, especially yeast cells, mammalian cells for example CHO cells or simian cells such as Vero cells, insect cells, plant cells, or protozoan cells.

For more details concerning the procedure to be applied in such a case, reference is made to the publication of Marks et al, J. Mol. Biol. 1991, 222:581-597.

Furthermore, starting from the immunoglobulins of the invention, or from fragments thereof, new immunoglobulins or derivatives can be prepared.

Accordingly immunoglobulins replying to the above given definitions can be prepared against determined antigens. Especially the invention provides monoclonal or polyclonal antibodies devoid of light polypeptide chains or antisera containing such antibodies and directed against determined antigens and for example against antigens of pathological agents such as bacteria, viruses or parasites. As example of antigens or antigenic determinants against which antibodies could be prepared, one can cite the envelope glycoproteins of viruses or peptides thereof, such as the external envelope glycoprotein of a HIV virus, the surface antigen of the hepatitis B virus.

Immunoglobulins of the invention can also be directed against a protein, hapten, carbohydrate or nucleic acid.

Particular antibodies according to the invention are directed against the galactosyl-1-3-galactose epitope.

The immunoglobulins of the invention allow further the preparation of combined products such as the combination of the heavy-chain immunoglobulin or a fragment thereof with a toxin, an enzyme, a drug, a hormone.

As example one can prepare the combination of a heavy-chain immunoglobulin bearing an antigen binding site recognizing a myeloma immunoglobulin epitope with the abrin or mistletoe lectin toxin. Such a construct would have its uses in patient specific therapy.

Another advantageous combination is that one can prepare between a heavy-chain immunoglobulins recognizing an insect gut antigen with a toxin specific for insects such as the toxins of the different serotypes of *Bacillus thuringiensis* or *Bacillus sphaericus*. Such a construct cloned into plants can be used to increase the specificity or the host range of existing bacterial toxins.

The invention also proposes antibodies having different specificities on each heavy polypeptide chains. These multifunctional, especially bifunctional antibodies could be prepared by combining two heavy chains of immunoglobulins of the invention or one heavy chain of an immunoglobulin of the invention with a fragment of a four-chain model immunoglobulin.

The invention also provides hetero-specific antibodies which can be used for the targeting of drugs or any biological substance like hormones. In particular they can be used to selectively target hormones or cytokines to a limited category of cells. Examples are a combination of a murine or human antibody raised against interleukin 2 (IL<sub>2</sub>) and a heavy-chain antibody raised against CD<sub>4</sub> cells. This could be used to reactivate CD<sub>4</sub> cells which have lost their IL<sub>2</sub> receptor.

The heavy-chain immunoglobulins of the invention can also be used for the preparation of hetero-specific antibodies. These can be achieved either according to the above described method by reduction of the bridges between the different chains and reoxidation, according to the usual techniques, of two antibodies having different specificities, but it can also be achieved by serial cloning of two antibodies for instance in the Immuno pBS vector.

In such a case, a first gene corresponding to the V<sub>H</sub>H domain comprised between Xho site and a Spe site is prepared as described above. A second gene is then prepared through an analogous way by using as 5' extremity a primer containing a Spe site, and as 3' extremity a primer containing a termination codon and an EcoRI site. The vector is then digested with EcoRI and XhoI and further both V<sub>H</sub>H genes are digested respectively by XhoI and by Spe/EcoRI.

After ligation, both immunoglobulin genes are serially cloned. The spacing between both genes can be increased by the introduction of addition codons within the 5' SpeI primer.

In a particular embodiment of the invention, the hinge region of IgG2 immunoglobulins according to the invention is semi-rigid and is thus appropriate for coupling proteins. In such an application proteins or peptides can be linked to various substances, especially to ligands through the hinge region used as spacer. Advantageously the fragment comprises at least 6 amino acids.

According to the invention it is interesting to use a sequence comprising a repeated sequence Pro-X, X being any amino-acid and preferably Gln, Lys or Glu, especially a fragment composed of at least a 3-fold repeat and preferably of a 12-fold repeat, for coupling proteins to ligand, or for assembling different protein domains.

The hinge region or a fragment thereof can also be used for coupling proteins to ligands or for assembling different protein domains.

Usual techniques for the coupling are appropriate and especially reference may be made to the technique of protein engineering by assembling cloned sequences.

The antibodies according to this invention could be used as reagents for the diagnosis in vitro or by imaging techniques. The immunoglobulins of the invention could be labelled with radio-isotopes, chemical or enzymatic markers or chemiluminescent markers.

As example and especially in the case of detection or observation with the immunoglobulins by imaging techniques, a label like technetium, especially technetium 99 is advantageous. This label can be used for direct labelling by a coupling procedure with the immunoglobulins or fragments thereof or for indirect labelling after a step of preparation of a complex with the technetium.

Other interesting radioactive labels are for instance indium and especially indium 111, or iodine, especially I<sup>131</sup>, I<sup>125</sup> and I<sup>123</sup>.

For the description of these techniques reference is made to the FR patent application published under number 2649488.

In these applications the small size of the V<sub>H</sub>H fragment is a definitive advantage for penetration into tissue.

The invention also concerns monoclonal antibodies reacting with anti-idiotypes of the above-described antibodies.

The invention also concerns cells or organisms in which heavy-chain immunoglobulins have been cloned. Such cells or organisms can be used for the purpose of producing heavy-chain immunoglobulins having a desired preselected specificity, or corresponding to a particular repertoire. They can also be produced for the purpose of modifying the metabolism of the cell which expresses them. In the case of modification of the metabolism of cells transformed with the sequences coding for heavy-chain immunoglobulins, these produced heavy-chain immunoglobulins are used like antisense DNA. Antisense DNA is usually involved in blocking the expression of certain genes such as for instance the variable surface antigen of trypanosomes or other pathogens. Likewise, the production or the activity of certain proteins or enzymes could be inhibited by expressing anti-antibodies against this protein or enzyme within the same cell.

The invention also relates to a modified 4-chain immunoglobulin or fragments thereof, the  $V_H$  regions of which have been partially replaced by specific sequences or amino acids of heavy chain immunoglobulins, especially by sequences of the  $V_{HH}$  domain. A particular modified  $V_H$  domain of a four-chain immunoglobulin, is characterized in that the leucine, proline or glutamine in position 45 of the  $V_H$  regions has been replaced by other amino acids and preferably by arginine, glutamic acid or cysteine.

A further modified  $V_H$  or  $V_L$  domain of a four-chain immunoglobulin, is characterized by linking of CDR loops together or to FW regions by the introduction of paired cysteines, the CDR region being selected between the CDR<sub>1</sub> and the CDR<sub>2</sub>, the FW region being the FW<sub>2</sub> region, and especially in which one of the cysteines introduced is in position 31, 33 of the CDR<sub>1</sub>, or 45 of FW<sub>2</sub>, and the other in CDR<sub>2</sub>.

Especially the introduction of paired cysteines is such that the CDR<sub>2</sub> loop is linked to the FW<sub>2</sub> or CDR<sub>1</sub> domain and more especially the cysteine of the CDR<sub>2</sub> of the  $V_H$  is linked to a cysteine in position 31, 33 of the CDR<sub>1</sub>, or in position 45 of FW<sub>2</sub>.

In another embodiment of the invention, plant cells can be modified by the heavy-chain immunoglobulins according to the invention, in order that they acquire new properties or increased properties.

The heavy-chain immunoglobulins of the invention can be used for gene therapy of cancer for instance by using antibodies directed against proteins present on the tumor cells.

In such a case, the expression of one or two  $V_{HH}$  genes can be obtained by using vectors derived from parvo or adeno viruses. The parvo viruses are characterized by the fact that they are devoid of pathogenicity or almost not pathogenic for normal human cells and by the fact that they are capable of easily multiplying in cancer cells (Russell S. J. 1990, Immunol. Today 11: 196-200).

The heavy-chain immunoglobulins are for instance cloned within HindIII/XbaI sites of the infectious plasmid of the murine MVM virus (pMVM984). (Merchinsky et al. 1983, J. Virol. 47, 227-232) and then placed under the control of the MVM38 promoter.

The gene of the  $V_{HH}$  domain is amplified by PCR by using a 5' primer containing an initiation codon and a HindIII site, the 3' primer containing a termination codon and a XbaI site.

This construct is then inserted between positions 2650 (HindIII) and 4067 (XbaI) within the plasmid.

The efficiency of the cloning can be checked by transfection. The vector containing the antibody is then introduced in permissive cells (NB-E) by transfection.

The cells are recovered after two days and the presence of  $V_{HH}$  regions is determined with an ELISA assay by using rabbit antiserum reacting with the  $V_{HH}$  part.

The invention further allows the preparation of catalytic antibodies through different ways. The production of antibodies directed against components mimicking activated states of substrates (as example vanadate as component mimicking the activated state of phosphate in order to produce their phosphatase activities, phosphonate as compound mimicking the peptidic binding in order to produce proteases) permits to obtain antibodies having a catalytic function. Another way to obtain such antibodies consists in performing a random mutagenesis in clones of antibodies for example by PCR, in introducing abnormal bases during the amplification of clones. These amplified fragments obtained by PCR are then introduced within an appropriate vector for cloning. Their expression at the surface of the bacteria permits the detection by the substrate of clones having the enzymatic activity. These two approaches can of course be combined. Finally, on the basis of the data available on the structure, for example the data obtained by X-ray crystallography or NMR, the modifications can be directed. These modifications can be performed by usual techniques of genetic engineering or by complete synthesis. One advantage of the  $V_{HH}$  of the heavy chain immunoglobulins of the invention is the fact that they are sufficiently soluble.

The heavy chain immunoglobulins of the invention can further be produced in plant cells, especially in transgenic plants. As example the heavy chain immunoglobulins can be produced in plants using the pMon530 plasmid (Roger et al. Meth Enzym 153 1566 1987) constitutive plant expression vector as has been described for classical four chain antibodies (Hiat et al. Nature 342 76-78, 1989) once again using the appropriate PCR primers as described above, to generate a DNA fragment in the right phase.

Other advantages and characteristics of the invention will become apparent in the examples and figures which follow.

## FIGURES

FIG. 1(A), (B), and (C): Characterisation and purification of camel IgG by affinity chromatography on Protein A and Protein G sepharose (Pharmacia)

Figure (A) shows, after reduction, the SDS-PAGE protein profile of the adsorbed and non adsorbed fractions of *Camelus dromedarius* serum. The fraction adsorbed on Protein A and eluted with NaCl 0.15M acetic acid 0.58% show upon reduction (lane c) three heavy chain components of respectively 50, 46 and 43 Kd and light chain (rabbit IgG in lane a). The fractions adsorbed on a Protein G Sepharose (Pharmacia) derivative which has been engineered to delete the albumin binding region (lane e) and eluted with 0.1M gly HCl pH 2.7 lacks the 40 Kd heavy chain which is recovered in the non adsorbed fraction (lane f). None of these components are present in the fraction non adsorbed on Protein A (lane d), lane b contains the molecular weight markers.

Figures (B) and (C) By differential elution, immunoglobulin fractions containing the 50 and 43 Kd heavy chain can be separated. 5 ml of *C. dromedarius* serum is adsorbed onto a 5 ml Protein G sepharose column and the column is extensively washed with 20 mM phosphate buffer, pH 7.0. Upon elution with pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid) a 100 Kd component is eluted which upon reduction yields a 43 Kd heavy chain, (lane 1). After column eluant absorbance has fallen to background level a second immunoglobulin component of 170 Kd can be eluted with pH 2.7 buffer (0.1M glycine HCl). This fraction upon reduction

yields a 50 Kd heavy chain and a board light chain band (lane 2). The fraction non adsorbed on Protein G is then brought on a 5 ml Protein A Sepharose column. After washing and elution with pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid) a third immunoglobulin of 100 Kd is obtained which consists solely of 46 Kd heavy chains (lane 3)

FIG. 2(A) and (B): Immunoglobulins of *Camelus bactrianus*, *Lama vicugna*, *Lama glama* and *Lama pacos* to Protein A (A lanes) and to Protein G (G lanes) analyzed on SDS-PAGE before Figure (A) and after reduction Figure (B)

10  $\mu$ l of serum obtained from the different species were added to Eppendorf<sup>®</sup> tubes containing 10 mg of Protein A or Protein G sepharose suspended in 400  $\mu$ l of pH 8.3 immunoprecipitation buffer (NaCl 0.2 M, Tris 0.01M; EDTA 0.01M, Triton X100 1%, ovalbumin 0.1%). The tubes were slowly rotated for 2 hours at 4° C. After centrifugation the pellets were washed 3 times in buffer and once in buffer in which the Triton and ovalbumin had been omitted. The pellets were then resuspended in the SDS-PAGE sample solution 70  $\mu$ l per pellet with or without dithiothreitol as reductant. After boiling for 3 min at 100° C., the tubes were centrifuged and the supernatants analysed. In all species examined the unreduced fractions Figure (A) contain in addition to molecules of approximately 170 Kd also smaller major components of approximately 100 Kd. In the reduced sample Figure (B) the constituent heavy and light chains are detected. In all species a heavy chain component (marked by an asterisk \*) is present in the material eluted from the Protein A but absent in the material eluted from the Protein G.

FIG. 3(A)(B)(C): IgG<sub>1</sub>, IgG<sub>2</sub> and IgG<sub>3</sub> were prepared from serum obtained from healthy or *Trypanosoma evansi* infected *Camelus dromedarius* (CAIT titer 1/160 (3) and analysed by radioimmuno-precipitation or Western Blotting for anti trypanosome activity

Figure 3<sup>55</sup> <sup>35</sup>S methionine labelled *Trypanosoma evansi* antigens lysate (500,000 counts) was added to Eppendorf tubes containing 10  $\mu$ l of serum or, 20  $\mu$ g of IgG<sub>1</sub>, IgG<sub>2</sub> or IgG<sub>3</sub> in 200  $\mu$ l of pH 8.3 immunoprecipitation buffer containing 0.1M TLCK as proteinase inhibitor and slowly rotated at 4° C. during one hour. The tubes were then supplemented with 10 mg of Protein A Sepharose suspended in 200  $\mu$ l of the same pH 8.3 buffer and incubated at 4° C. for an additional hour. After washing and centrifugation at 15000 rpm for 12 s, each pellet was resuspended in 75  $\mu$ l SDS-PAGE sample solution containing DTT and heated for 3 min. at 100° C. After centrifugation in an Eppendorf microfuge at 15000 rpm for 30 s, 5  $\mu$ l of the supernatant was saved for radioactivity determination and the remainder analysed by SDS-PAGE and fluorography. The counts/5  $\mu$ l sample are inscribed on for each line.

Figure (B) (C) 20  $\mu$ g of IgG<sub>1</sub>, IgG<sub>2</sub> and IgG<sub>3</sub> from healthy and trypanosome infected animals were separated by SDS-PAGE without prior reduction or heating. The separated samples were then electro transferred to a nitrocellulose membrane, one part of the membrane was stained with Ponceau Red to localise the protein material and the remainder incubated with 1% ovalbumin in TST buffer (Tris 10 mM, NaCl 150 mM, Tween 0.05%) to block protein binding sites. After blocking, the membrane was extensively washed with TST buffer and incubated for 2 hours with <sup>35</sup>S-labelled trypanosome antigen. After extensive washing, the membrane was dried and analysed by autoradiography. To avoid background and unspecific binding, the labelled trypanosome lysate was filtered through a 45 $\mu$  millipore filter and incubated with healthy camel immunoglobulin and ovalbumin adsorbed on a nitrocellulose membrane.

FIG. 4(A) and (B): Purified IgG3 of the camel, by affinity chromatography on Protein A Sepharose are partially digested with papain and separated on Protein A sepharose.

14 mg of purified IgG3 were dissolved in 0.1M phosphate buffer pH 7.0 containing 2 mM EDTA. They were digested by 1 hour incubation at 37° C. with mercurypapain (1% enzyme to protein ratio) activated by 5.10<sup>-3</sup>M cysteine. The digestion was blocked by the addition of excess iodoacetamide (4.10<sup>-3</sup>M) (13). After centrifugation of the digest in an ependorf centrifuge for 5min at 15000 rpm, the papain fragments were separated on a protein A Sepharose column into binding (B) and non binding (NB) fractions. The binding fraction was eluted from the column with 0.1M glycine HCl buffer pH 1.7.

FIG. 5: Schematic presentation of a model for IgG3 molecules devoid of light chains.

FIG. 6: Schematic representation of immunoglobulins having heavy polypeptide chains and devoid of light chains, regarding conventional four-chain model immunoglobulin. Representation of a hinge region.

FIG. 7: Alignment of 17 V<sub>H</sub>H DNA sequences of Camel heavy chain immunoglobulins SEQ ID NOS:92-108.

FIG. 8: Expression and purification of the camel V<sub>H</sub>H21 protein from *E. coli*

# I HEAVY CHAIN ANTIBODIES IN CAMELIDS

When *Camelus dromedarius* serum is adsorbed on Protein G sepharose, an appreciable amount (25-35%) of immunoglobulins (Ig) remains in solution which can then be recovered by affinity chromatography on Protein A sepharose (FIG. 1A). The fraction adsorbed on Protein G can be differentially eluted into a tightly bound fraction (25%) consisting of molecules of an unreduced apparent molecular weight (MW) of 170 Kd and a more weakly bound fraction (30-45%) having an apparent molecular weight of 100 Kd (FIG. 1B). The 170 Kd component when reduced yields 50 Kd heavy chains and large 30 Kd light chains. The 100 Kd fraction is totally devoid of light chains and appears to be solely composed of heavy chains which after reduction have an apparent MW of 43 Kd (FIG. 1C). The fraction which does not bind to Protein G can be affinity purified and eluted from a Protein A column as a second 100 Kd component which after reduction appears to be composed solely of 46 Kd heavy chains.

The heavy chain immunoglobulins devoid of light chains total up to 75% of the molecules binding to Protein A.

As all three immunoglobulins bind to Protein A we refer to them as IgG: namely IgG<sub>1</sub> (light chain and heavy chain  $\gamma$ 1 (50 Kd) binding to Protein G and IgG<sub>2</sub> (heavy chain  $\gamma$ 2 (46 Kd) non binding to Protein G and IgG<sub>3</sub> (heavy chain  $\gamma$ 3 (43 Kd) binding to Protein G. There is a possibility that these three sub(classes) can be further subdivided.

A comparative study of old world camelids (*Camelus bactrianus* and *Camelus dromedarius*) and new world camelids (*Lama pacos*, *Lama glama*, *Lama vicugna*) showed that heavy chain immunoglobulins are found in all species examined, albeit with minor differences in apparent molecular weight and proportion. The new world camelids differs from the old world camelids in having a larger IgG<sub>3</sub> molecule (heavy chain immunoglobulin binding to Protein G) in which the constituent heavy chains have an apparent molecular weight of 47 Kd (FIG. 2A and B).

The abundance of the heavy chain immunoglobulins in the serum of camelids raises the question of what their role is in the immune response and in particular whether they



bear antigen binding specificity and if so how extensive is the repertoire. This question could be answered by examining the immunoglobulins from *Trypanosoma evansi* infected camels (*Camelus dromedarius*).

For this purpose, the corresponding fractions of IgG<sub>1</sub>, IgG<sub>2</sub>, IgG<sub>3</sub> were prepared from the serum of a healthy camel and from the serum of camels with a high antityranosome titer, measured by the Card Agglutination Test (3). In radioimmunoprecipitation, IgG<sub>1</sub>, IgG<sub>2</sub> and IgG<sub>3</sub> derived from infected camel indicating extensive repertoire heterogeneity and complexity (FIG. 3A) were shown to bind a large number of antigens present in a <sup>35</sup>S methionine labelled trypanosome lysate.

In blotting experiments <sup>35</sup>S methionine labelled trypanosome lysate binds to SDS PAGE separated IgG<sub>1</sub>, IgG<sub>2</sub> and IgG<sub>3</sub> obtained from infected animals (FIG. 3B and C).

This leads us to conclude that the camelid heavy chain IgG<sub>2</sub> and IgG<sub>3</sub> are bona fide antigen binding antibodies.

An immunological paradigm states that an extensive antibody repertoire is generated by the combination of the light and heavy chain variable V region repertoires (6). The heavy chain immunoglobulins of the camel seem to contradict this paradigm.

Immunoglobulins are characterized by a complex I.E.F. (isoelectric focussing) pattern reflecting their extreme heterogeneity. To determine whether the two heavy chains constituting the IgG<sub>2</sub> and IgG<sub>3</sub> are identical or not, the isoelectric focussing (I.E.F.) pattern were observed before and after chain separation by reduction and alkylation using iodoacetamide as alkylating agent.

As this alkylating agent does not introduce additional charges in the molecule, the monomers resulting from the reduction and alkylation of a heavy chain homodimer will have practically the same isoelectric point as the dimer, whereas if they are derived from a heavy chain heterodimer, the monomers will in most cases differ sufficiently in isoelectric point to generate a different pattern in I.E.F.

Upon reduction, and alkylation by iodoacetamide the observed pattern is not modified for the *Camelus dromedarius* IgG<sub>2</sub> and IgG<sub>3</sub> indicating that these molecules are each composed of two identical heavy chains which migrate to the same position as the unreduced molecule they originated from.

In contrast, the I.E.F. pattern of IgG<sub>1</sub> is completely modified after reduction as the isoelectric point of each molecule is determined by the combination of the isoelectric points of the light and heavy chains which after separation will each migrate to a different position.

These findings indicate that the heavy chains alone can generate an extensive repertoire and question the contribution of the light chain to the useful antibody repertoire. If this necessity be negated, what other role does the light chain play.

Normally, isolated heavy chain from mammalian immunoglobulins tend to aggregate considerably but are only solubilized by light chains (8, 9) which bind to the C<sub>H</sub>1 domain of the heavy chain.

In humans and in mice a number of spontaneous or induced myelomas produce a pathological immunoglobulin solely composed of heavy chains (heavy chain disease). These myeloma protein heavy chains carry deletions in the C<sub>H</sub>1 and V<sub>H</sub>H domains (10). The reason why full length heavy chains do not give rise to secreted heavy chain in such pathological immunoglobulins seems to stem from the fact that the synthesis of Ig involves a chaperoning protein, the

immunoglobulin heavy chain binding protein or BiP (11), which normally is replaced by the light chain (12). It is possible that the primordial role of the light chain in the four-chain model immunoglobulins is that of a committed heavy chain chaperon and that the emergence of light chain repertoires has just been an evolutionary bonus.

The camelid  $\gamma$ 2 and  $\gamma$ 3 chains are considerably shorter than the normal mammalian  $\gamma$  chain. This would suggest that deletions have occurred in the C<sub>H</sub>1 domain. Differences in sizes of the  $\gamma$ 2 and  $\gamma$ 3 immunoglobulins of old and new world camelids suggests that deletions occurred in several evolutionary steps especially in the C<sub>H</sub>1 domain.

## II THE HEAVY CHAIN IMMUNOGLOBULINS OF THE CAMELIDS LACK THE C<sub>H</sub>1 DOMAIN.

The strategy followed for investigating the heavy chain immunoglobulin primary structure is a combination of protein and cDNA sequencing; the protein sequencing is necessary to identify sequence stretches characteristic of each immunoglobulin. The N-terminal of the immunoglobulin being derived from the heavy chain variable region repertoire only yields information on the V<sub>H</sub>H subgroups (variable region of the heavy chain) and cannot be used for class or subclass identification. This means that sequence data had to be obtained from internal enzymatic or chemical cleavage sites.

A combination of papain digestion and Protein A affinity chromatography allowed the separation of various fragments yielding information on the general structure of IgG3.

The IgG3 of the camel (*Camelus dromedarius*) purified by affinity chromatography on Protein A Sepharose were partially digested with papain and the digest was separated on Protein A Sepharose into binding and non binding fractions. These fractions were analysed by SDS PAGE under reducing and non reducing conditions (FIG. 4A and B).

The bound fraction contained two components, one of 28 Kd and one of 14.4 Kd, in addition to uncleaved or partially cleaved material. They were well separated by gel electrophoresis (from preparative 19% SDS-PAGE gels) under non reducing conditions and were further purified by electroelution (in 50 mM ammonium bicarbonate, 0.1% (w/v) SDS using a BioRad electro-eluter). After lyophilization of these electroeluted fractions, the remaining SDS was eliminated by precipitating the protein by the addition of 90% ethanol, mixing and incubating the mixture overnight at -20°C (14). The precipitated protein was collected in a pellet by centrifuging (15000 rpm, 5 min) and was used for protein sequencing. N-terminal sequencing was performed using the automated Edman chemistry of an Applied Biosystem 477A pulsed liquid protein sequencer. Amino acids were identified as their phenylthiohydantoin (PTH) derivatives using an Applied Biosystem 120 PTH analyser. All chemical and reagents were purchased from Applied Biosystems. Analysis of the chromatographic data was performed using Applied Biosystems software version 1.61. In every case the computer aided sequence analysis was confirmed by direct inspection of the chromatograms from the PTH analyser. Samples for protein sequencing were dissolved in either 50% (v/v) trifluoroacetic acid (TFA) (28Kd fragment) or 100% TFA (14Kd fragment). Samples of dissolved protein equivalent to 2000 pmol (28 Kd fragment) or 500 pmol (14Kd fragment) were applied to TFA-treated glass fibre discs. The glass fibre discs were coated with BioBrene (3 mg) and pre-cycled once before use.

N-terminal sequencing of the 28 Kd fragment yields a sequence homologous to the N-terminal part of  $\gamma$  C<sub>H</sub>2

domain and hence to the N-terminal end of the Fc fragment. The N-terminal sequence of the 14.4 Kd fragment corresponds to the last lysine of a  $\gamma C_H2$  and the N-terminal end of a  $\gamma C_H3$  domain (Table 1). The molecular weight (MW) of the papain fragments and the identification of their N-terminal sequences led us to conclude that the  $C_H2$  and  $C_H3$  domains of the  $\gamma 3$  heavy chains are normal in size and that the deletion must occur either in the  $C_H1$  or in the  $V_{HH}$  domain to generate the shorter  $\gamma 3$  chain. The fractions which do not bind to Protein A Sepharose contain two bands of 34 and 17 Kd which are more diffuse is SDS PAGE indicating that they originate from the variable N-terminal part of the molecule (FIG. 4A and B).

Upon reduction, a single diffuse band of 17 Kd is found indicating that the 34 Kd is a disulfide bonded dimer of the 17 Kd component. The 34 Kd fragment apparently contains the hinge and the N-terminal domain  $V_{HH}$ . The protein sequence data can be used to construct degenerate oligonucleotide primers allowing PCR amplification of cDNA or genomic DNA.

It has been shown that the cells from camel spleen imprint cells reacted with rabbit and anti camel immunoglobulin sera and that the spleen was hence a site of synthesis of at least one immunoglobulin class. cDNA was therefore synthesized from camel spleen mRNA. The conditions for the isolation of RNA were the following: total RNA was isolated from the dromedary spleen by the guanidium isothiocyanate method (15). mRNA was purified with oligo T-paramagnetic beads. cDNA synthesis is obtained using 1  $\mu$ g mRNA template, an oligodT primer and reverse transcriptase (BOEHRINGER MAN). Second strand cDNA is obtained using RNase H and *E. coli* DNA polymerase I according to the condition given by the supplier. Relevant sequences were amplified by PCR: 5 ng of cDNA was amplified by PCR in a 100  $\mu$ l reaction mixture (10 mM Tris-HCl pH 8.3, 50 mM KCl, 15 mM MgCl<sub>2</sub>, 0.01% (w/v) gelatin, 200  $\mu$ M of each dNTP and 25 pmoles of each primer) overlaid with mineral oil (Sigma). Degenerate primers containing EcoRI and KpnI sites and further cloned into pUC 18. After a round of denaturing and annealing (94° C. for 5 min and 54° C. for 5 min), 2 units of Taq DNA polymerase were added to the reaction mixture before subjecting it to 35 cycles of amplification: 1 min at 94° C. (denature) 1 min at 54° C. (anneal), 2 min at 72° C. (elongate). To amplify DNA sequences between  $V_{HH}$  and  $C_H2$  domains, (#72 clones), the PCR was performed in the same conditions with the exception that the annealing temperature was increased to 60° C.

One clone examined (#56/36) had a sequence corresponding to the N-terminal part of a  $C_H2$  domain identical to the sequence of the 28 Kd fragment. The availability of this sequence data allowed the construction of an exact 3' primer and the cloning of the region between the N-terminal end of the  $V_{HH}$  and the  $C_H2$  domain.

5' primers corresponding to the mouse  $V_{HH}$  (16) and containing a XhoI restriction site were used in conjunction with the 3' primer in which a KpnI site had been inserted and the amplified sequences were cloned into pBluescript<sup>®</sup>. Clone #56/36 which displayed two internal HaeIII sites was digested with this enzyme to produce a probe to identify PCR positive clones.

After amplification the PCR products were checked on a 1.2% (w/v) agarose gel. Cleaning up of the PCR products included a phenol-chloroform extraction followed by further purification by HPLC (GEN-PAC FAX column, Waters) and finally by using the MERMAID or GENE/CLEAN II kit, BIO 101, Inc) as appropriate. After these purification steps,

the amplified cDNA was then digested with EcoRI and KpnI for series #56 clones and with XhoI and KpnI for series #72 clones. A final phenol-chloroform extraction preceded the ligation into pUC 18 (series #56 clones) or into pBluescript<sup>®</sup> (series #72 clones).

All the clones obtained were smaller than the 860 base pairs to be expected if they possessed a complete  $V_{HH}$  and  $C_H1$  region. Partial sequence data corresponding to the N-terminal of the  $V_{HH}$  region reveals that out of 20 clones, 3 were identical and possibly not independent. The sequences obtained resemble the human subgroup III and the murine subgroups IIIa and IIIb (Table 2).

Clones corresponding to two different sets of  $C_H2$  protein sequences were obtained. A first set of sequences (#72/41) had a N-terminal  $C_H2$  region identical to the one obtained by protein sequencing of the 28 Kd papain fragments of the  $\gamma 3$  heavy chain, a short hinge region containing 3 cysteines and a variable region corresponding to the framework (FR4) residues encoded by the J minigenes adjoining the hinge. The  $C_H1$  domain is entirely lacking. This cDNA corresponds to the  $\gamma 3$  chain (Table 4).

In one closely related sequence (#72/1) the proline in position 259 is replaced by threonine.

The sequence corresponding to the  $C_H3$  and the remaining part of the  $C_H2$  was obtained by PCR of the cDNA using as KpnI primer a poly T in which a KpnI restriction site had been inserted at the 5' end. The total sequence of the  $\gamma 3$  chain corresponds to a molecular weight (MW) which is in good agreement with the data obtained from SDS PAGE electrophoresis.

The sequence of this  $\gamma 3$  chain presents similarities with other 7 chains except that it lacks the  $C_H1$  domain, the  $V_{HH}$  domain being adjacent to the hinge.

One or all three of the cysteines could be probably responsible for holding the two  $\gamma 3$  chains together.

These results have allowed us to define a model for the IgG3 molecule based on sequence and papain cleavage (FIG. 5).

Papain can cleave the molecule on each side of the hinge disulfides and also between  $C_H2$  and  $C_H3$ . Under non reducing conditions the  $V_{HH}$  domains of IgG3 can be isolated as disulfide linked dimer or as monomer depending on the site of papain cleavage.

A second set of clones #72/29 had a slightly different sequence for the  $C_H2$  and was characterized by a very long hinge immediately preceded by the variable domain. This hinge region has 3 cysteines at its C-terminal end in a sequence homologous to the  $\gamma 3$  hinge. Such second set of clones could represent the IgG2 subclass. For the constant part of the  $\gamma 3$  and also for the putative  $\gamma 2$ , most clones are identical showing the  $\gamma 2$  or  $\gamma 3$  specific sequences. A few clones such as #72/1 however show minor differences. For instance in the case of clones #72/1 two nucleotide differences are detected.

Several  $V_{HH}$  regions cDNA's have now been totally or partially sequenced with the exception of a short stretch at the N-terminal end which is primer derived.

Upon translation the majority shows the characteristic heavy chain Ser<sub>21</sub>, Cys<sub>22</sub> and Tyr<sub>90</sub> Tyr<sub>91</sub>, Cys<sub>92</sub> sequences, of the intra  $V_{HH}$  region disulfide bridge linking residues 22 and 92. All these clones have a sequence corresponding to the framework 4 (FR4) residues of the variable region immediately preceding the postulated hinge sequence (Table 3). This sequence is generated by the J minigenes and is in the majority of cases similar to the sequence encoded by the

human and murine J minigenes. The sequence length between region Cys<sub>82</sub> and the C-terminal end of the V<sub>H</sub>H regions is variable and, in the sequences determined, range from 25 to 37 amino-acids as one might expect from the rearrangements of J and D minigenes varying in length.

Several important questions are raised by the sole existence of these heavy chain immunoglobulins in a non pathological situation. First of all, are they bonafide antibodies? The heavy chain immunoglobulins obtained from trypanosome infected camels react with a large number of parasite antigens as shown in part of these examples. This implies that the camelid immune system generates an extensive number of binding sites composed of single V<sub>H</sub>H domains. This is confirmed by the diversity of the V<sub>H</sub>H regions of the heavy chain immunoglobulins obtained by PCR.

The second question is "how are they secreted?". The secretion of immunoglobulin heavy chains composing four-chain model immunoglobulins does not occur under normal conditions. A chaperoning protein, the heavy chain binding protein, or BIP protein, prevents heavy chains from being secreted. It is only when the light chain displaces the BIP protein in the endoplasmic reticulum that secretion can occur (13).

The heavy chain dimer found in the serum of human or mice with the so-called "heavy chain disease" lack the C<sub>H</sub>1 domains thought to harbour the BIP site (14). In the absence of the light chain the BIP protein can no longer bind and prevent the transport of the heavy chains.

The presence in camels of a IgG1 class composed of heavy and light chains making up between 25% and 50% of the total IgG molecules also raises the problem as to how maturation and class switching occurs and what the role of the light chain is. The camelid light chain appears unusually large and heterogeneous when examined in SDS PAGE.

The largest dimension of an isolated domain is 40 Å and the maximum attainable span between binding sites of a conventional IgG with C<sub>H</sub>1 and V<sub>H</sub>H will be of the order of 160 Å (2V<sub>H</sub>H+2C<sub>H</sub>1) (19). The deletion of C<sub>H</sub>1 domain in the two types of heavy chain antibodies devoid of light chains, already sequenced has, as a result, a modification of this maximum span (FIG. 6). In the IgG3 the extreme distance between the extremities of the V<sub>H</sub>H regions will be of the order of 80 Å (2V<sub>H</sub>H). This could be a severe limitation for agglutinating or cross linking. In the IgG2 this is compensated by the extremely long stretch of hinge, composed of a 12-fold repeat of the sequence Pro-X (where X is Gln, Lys or Glu) and located N-terminal to the hinge disulfide bridges. In contrast, in the human IgG3, the very long hinge which also apparently arose as the result of sequence duplication does not contribute to increase the distance spanning the two binding sites as this hinge is inter-spaced with disulfide bridges.

The single V<sub>H</sub>H domain could also probably allow considerably rotational freedom of the binding site versus the Fc domain.

Unlike myeloma heavy chains which result probably from C<sub>H</sub>1 deletion in a single antibody producing cell, or heavy chain antibodies produced by expression cloning (15); the camelid heavy chain antibodies (devoid of light chains) have emerged in a normal immunological environment and it is expected that they will have undergone the selective refinement in specificity and affinity accompanying B cell maturation.

#### Expression and Purification of the Camel V<sub>H</sub>H21 (DR21 on FIG. 7) Protein from *E. coli*

The clones can be expressed in several types of expression vectors. As an example using a commercially available vector Immuno PBS (Huse et al. Science (1989) 246, 1275), clones produced in Bluescript® according to the above described procedure, have been recovered by PCR using the same XhoI containing 5' primer and a new 3' primer, corresponding to residues 113-103 in the framework of the immunoglobulins, in which an Spe site has been constructed: TC TTA ACT AGT GAG GAG ACG GTG ACC TG SEQ ID NO: This procedure allowed the cloning of the V<sub>H</sub>H in the Xho/Spe site of the Immuno PBS vector. However, the 3' end of the gene was not in phase with the identification "tag" and the stop codon of the vector. To achieve this, the construct was cut with Spe and the 4 base overhangs were filled in, using the Klenow fragment after which the vector was religated.

The expression vector plasmid ipBS (immunopBS) (Stratagene) contains a pel B leader sequence which is used for immunoglobulin chain expression in *E. coli* under the promoter pLAC control, a ribosome binding site, and stop codons. In addition, it contains a sequence for a c-terminal decapeptide tag.

*E. coli* JM101 harboring the ipBS-V<sub>H</sub>H21 plasmid was grown in 1 l of TB medium with 100 µg/ml ampicillin and 0.1% glucose at 32° C. Expression was induced by the addition of 1 mM IPTG (final concentration) at an OD<sub>260</sub> of 1.0. After overnight induction at 28° C., the cells were harvested by centrifugation at 4,000 g for 10 min (4° C.) and resuspended in 10 ml TES buffer (0.2M Tris-HCl pH 8.0, 0.5 mM EDTA, 0.5M sucrose). The suspension was kept on ice for 2 hours. Periplasmic proteins were removed by osmotic shock by addition of 20 ml TES buffer diluted 1:4 v/v with water, kept on ice for one hour and subsequently centrifuged at 12,000 g for 30 min. at 4° C. The supernatant periplasmic fraction was dialysed against Tris-HCl pH 8.8, NaCl 50 mM, applied on a fast Q Sepharose flow (Pharmacia) column, washed with the above buffer prior and eluted with a linear gradient of 50 mM to 1M NaCl in buffer. Fractions containing the V<sub>H</sub>H protein were further purified on a Superdex 75 column (Pharmacia) equilibrated with PBS buffer (0.01M phosphate pH 7.2, 0.15M NaCl). The yield of purified V<sub>H</sub>H protein varies from 2 to 5 mg/l cell culture.

Fractions were analyzed by SDS-PAGE(I). Positive identification of the camel V<sub>H</sub>H antibody fragment was done by Western Blot analysis using antibody raised in rabbits against purified camel IgG<sub>H</sub>3 and an anti-rabbit IgG-alkaline phosphatase conjugate (II).

As protein standards (Pharmacia) periplasmic proteins prepared from 1 ml of IPTG-induced JM101/ipBS V<sub>H</sub>H21 were used. FIGS. 8A and 8B shows: C,D: fractions from fast S Sepharose column chromatography (C: eluted at 650 mM NaCl D: eluted at 700 mM NaCl) E,F: fractions from Superdex 75 column chromatography.

As can be seen, the major impurity is eliminated by ionexchange chromatography and the bulk of the remaining impurities are eliminated by gel filtration.



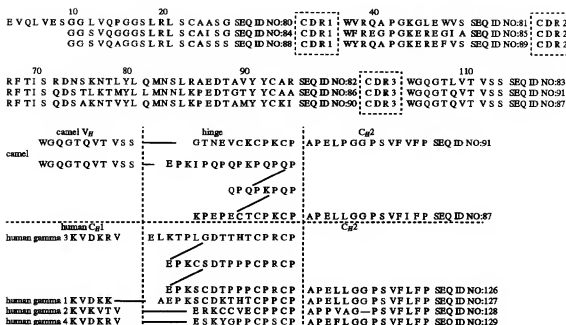
TABLE 3-continued

Comparison of some Framework 4 residues found in the Camel V<sub>H</sub>H1 region with the Framework 4 residues corresponding to the consensus regions of the Human and Mouse J regions.

Frame Work 4									
W	G	Q	G	T	Q	V	T	V	S
W	G	Q	G	T	L	V	T	V	S
W	G	R	G	T	Q	V	T	V	S
W	G	Q	G	T	H	V	T	V	S
W	G	Q	G	I	Q	V	T	A	S
								SEQ ID NO:8	# 72/19 - # 72/3
								SEQ ID NO:9	1 Clone
								SEQ ID NO:59	# 72/24
								SEQ ID NO:60	# 72/21
								SEQ ID NO:61	# 72/16

TABLE 4

SEQ ID NO: 125										MOUSE	V <sub>H</sub> H	Sequence																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
95	d	y	y	g	s	s	-----	y	---	f	-----	d	v	W	G	A	G	T	T	V	T	V	S	S																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			</



## REFERENCES

- Ward, E. S., Gussow, D., Griffiths, A. D., Jones, P. T. and Winter G. Nature 341, 544-546 (1989).
- Ungar-Waron H., Ellise E., Gluckman A. and Trainin Z. Isr. J. Vet. Med., 43, 198-203 (1987).
- Bajanya Songa E. and Hamers R., Ann. Soc. Belge Med. trop., 68, 233-240 (1988).
- Edelman G. M., Olins D. E., Gally J. A. and Zinder N. D., Proc. Nat. Acad. Sc., 50, 753 (1963).
- Franeck F. and Nezin R. S., Biokhimiya, 28, 193 (1963).
- Roitt I. M., Brostoff J. and Male D. K., Immunology, Gower Med. Pub. London. New-York, p. 92. (1985).
- Schiffer M., Girdling R. L., Ely K. R. and Edmundson B., Biochemistry, 12, 4620-4631 (1973).
- Fleischman J. B., Pain R. H. and Porter R. R., Arch. Biochem. Biophys., Suppl. 1, 174 (1962).
- Roholt O., Onoue K. and Pressman D., PNAS 51, 173-178 (1964).
- Seligmann M., Mihaesco E., Freud'homme J. L., Danon F. and Brouet J. C., Immunological Rev., 48, 145-167 (1979).
- Hendershot L., Bole D., Kohler G. and Kearney J. F., The Journal of Cell Biology, 104, 761-767 (1987).
- Hendershot L. M., The Journal of Cell Biology, 111, 829-837 (1990).
- Hamers-Casterman, C., E. Wittout, W. Van der Loo and R. Hamers, Journal of Immunogenetics, 6, 373-381 (1979).
- Applied Biosystems—Ethanol Precipitation of Electro Eluted Electrolysed Sample. Issue no. 27.
- Maniatis, T. E. F. Fritsch and J. Sambrook, Molecular Cloning. A Laboratory Manual (1988).
- Sastry et al., PNAS, 86, 5728, (1989).
- Sanger, F., S. Nicklen and A. R. Coulson, Proc. Natl. Acad. Sci., U.S.A., 74, 5463-5467 (1977).
- Kabat E. A., Tai Te Wu, M. Reid-Miller, H. M. Perry and K. S. Gottesman, U.S. Dpt of Health and Human Services, Public Health Service, National Institutes of Health (1987).
- Valentine, R. C. and N. M. Geen, J. M. B., 27, 615-617 (1967).

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(1.1) NUMBER OF SEQUENCES: 130

## (2) INFORMATION FOR SEQ ID NO:1:

## (2.1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(2.2) MOLECULE TYPE: peptide

(ix) FEATURE:  
 (A) NAME/KEY: Domain  
 (B) LOCATION: 1..22  
 (D) OTHER INFORMATION: /label=FRAMEWORK 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1

```
Gly  Gly  Ser  Val  Gln  Thr  Gly  Gly  Ser  Leu  Arg  Leu  Ser  Cys  Glu  Ile
 1              5              10              15
Ser  Gly  Leu  Thr  Phe  Asp
                20
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
 (A) NAME/KEY: Domain  
 (B) LOCATION: 1..22  
 (D) OTHER INFORMATION: /label=FRAMEWORK 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2

```
Gly  Gly  Ser  Val  Gln  Thr  Gly  Gly  Ser  Leu  Arg  Leu  Ser  Cys  Ala  Val
 1              5              10              15
Ser  Gly  Phe  Ser  Phe  Ser
                20
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
 (A) NAME/KEY: Domain  
 (B) LOCATION: 1..22  
 (D) OTHER INFORMATION: /label=FRAMEWORK 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3

```
Gly  Gly  Ser  Glu  Gln  Gly  Gly  Gly  Ser  Leu  Arg  Leu  Ser  Cys  Ala  Ile
 1              5              10              15
Ser  Gly  Tyr  Thr  Tyr  Gly
                20
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
 (A) NAME/KEY: Domain  
 (B) LOCATION: 1..22  
 (D) OTHER INFORMATION: /label=FRAMEWORK 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4

-continued

---

Gly	Gly	Ser	Val	Gln	Pro	Gly	Gly	Ser	Leu	Thr	Leu	Ser	Cys	Thr	Val
1				5					10					15	
Ser	Gly	Ala	Thr	Tyr	Ser										
					20										

( 2 ) INFORMATION FOR SEQ ID NO:5:

```

(i ) SEQUENCE CHARACTERISTICS:
      (A ) LENGTH: 22 amino acids
      (B ) TYPE: amino acid
      (C ) STRANDEDNESS: single
      (D ) TOPOLOGY: linear

(ii ) MOLECULE TYPE: peptide

(ix ) FEATURE:
      (A ) NAME/KEY: Domain
      (B ) LOCATION: 1..22
      (D ) OTHER INFORMATION: /label=FRAMEWORK 1

(xi ) SEQUENCE DESCRIPTION: SEQ ID NO:5:
      Gly  Gly  Ser  Val  Gln  Ala  Gly  Gly  Ser  Leu  Arg  Leu  Ser  Cys  Thr  Gly
       1           5           10           15
      Ser  Gly  Phe  Pro  Tyr  Ser
       20

```

( 2 ) INFORMATION FOR SEQ ID NO:6:

```

(i ) SEQUENCE CHARACTERISTICS:
      (A ) LENGTH: 21 amino acids
      (B ) TYPE: amino acid
      (C ) STRANDEDNESS: single
      (D ) TOPOLOGY: linear

(ii ) MOLECULE TYPE: peptide

(ix ) FEATURE:
      (A ) NAME/KEY: Domain
      (B ) LOCATION: 1..21
      (D ) OTHER INFORMATION: /label=FRAMEWORK 1

(xi ) SEQUENCE DESCRIPTION: SEQ ID NO:6:
      Gly  Gly  Ser  Val  Gln  Ala  Gly  Gly  Ser  Leu  Arg  Leu  Ser  Cys  Val  Ala
       1           5           10           15
      Gly  Phe  Gly  Thr  Ser
       20

```

( 2 ) INFORMATION FOR SEQ ID NO:7:

```

(i ) SEQUENCE CHARACTERISTICS:
      (A ) LENGTH: 21 amino acids
      (B ) TYPE: amino acid
      (C ) STRANDEDNESS: single
      (D ) TOPOLOGY: linear

(ii ) MOLECULE TYPE: peptide

(ix ) FEATURE:
      (A ) NAME/KEY: Domain
      (B ) LOCATION: 1..21
      (D ) OTHER INFORMATION: /label=FRAMEWORK 1

(xi ) SEQUENCE DESCRIPTION: SEQ ID NO:7:
      Gly  Gly  Ser  Val  Gln  Ala  Gly  Gly  Ser  Leu  Arg  Leu  Ser  Cys  Val  Ser
       1           5           10           15
      Phe  Ser  Pro  Ser  Ser
       20

```

( 2 ) INFORMATION FOR SEQ ID NO:8:



-continued

---

```

( i ) SEQUENCE CHARACTERISTICS:
  ( A ) LENGTH: 11 amino acids
  ( B ) TYPE: amino acid
  ( C ) STRANDEDNESS: single
  ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
  ( A ) NAME/KEY: Domain
  ( B ) LOCATION: 1..11
  ( D ) OTHER INFORMATION: /label=FRAMEWORK 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:8:

  Trp  Gly  Gln  Gly  Thr  Gln  Val  Thr  Val  Ser  Ser
   1          5          10

( 2 ) INFORMATION FOR SEQ ID NO:9:

( i ) SEQUENCE CHARACTERISTICS:
  ( A ) LENGTH: 11 amino acids
  ( B ) TYPE: amino acid
  ( C ) STRANDEDNESS: single
  ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
  ( A ) NAME/KEY: Domain
  ( B ) LOCATION: 1..11
  ( D ) OTHER INFORMATION: /label=FRAMEWORK 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:9:

  Trp  Gly  Gln  Gly  Thr  Leu  Val  Thr  Val  Ser  Ser
   1          5          10

( 2 ) INFORMATION FOR SEQ ID NO:10:

( i ) SEQUENCE CHARACTERISTICS:
  ( A ) LENGTH: 11 amino acids
  ( B ) TYPE: amino acid
  ( C ) STRANDEDNESS: single
  ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
  ( A ) NAME/KEY: Domain
  ( B ) LOCATION: 1..11
  ( D ) OTHER INFORMATION: /label=FRAMEWORK 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:10:

  Trp  Gly  Gln  Gly  Ala  Gln  Val  Thr  Val  Ser  Ser
   1          5          10

( 2 ) INFORMATION FOR SEQ ID NO:11:

( i ) SEQUENCE CHARACTERISTICS:
  ( A ) LENGTH: 11 amino acids
  ( B ) TYPE: amino acid
  ( C ) STRANDEDNESS: single
  ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
  ( A ) NAME/KEY: Domain
  ( B ) LOCATION: 1..11
  ( D ) OTHER INFORMATION: /label=FRAMEWORK 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:11:

  Trp  Gly  Gln  Gly  Thr  Gln  Val  Thr  Ala  Ser  Ser

```

-continued

1	5	10
( 2 ) INFORMATION FOR SEQ ID NO:12:		
( i ) SEQUENCE CHARACTERISTICS:		
( A ) LENGTH: 11 amino acids		
( B ) TYPE: amino acid		
( C ) STRANDEDNESS: single		
( D ) TOPOLOGY: linear		
( i i ) MOLECULE TYPE: peptide		
( i x ) FEATURE:		
( A ) NAME/KEY: Domain		
( B ) LOCATION: 1..11		
( D ) OTHER INFORMATION: /label=FRAMEWORK 4		
( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:12:		
Arg Gly Glu Gly Thr	Gln Val Thr Val Ser Leu	
1	5	10
( 2 ) INFORMATION FOR SEQ ID NO:13:		
( i ) SEQUENCE CHARACTERISTICS:		
( A ) LENGTH: 25 amino acids		
( B ) TYPE: amino acid		
( C ) STRANDEDNESS: single		
( D ) TOPOLOGY: linear		
( i i ) MOLECULE TYPE: peptide		
( i x ) FEATURE:		
( A ) NAME/KEY: Domain		
( B ) LOCATION: 1..14		
( D ) OTHER INFORMATION: /label=CDR3		
( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:13:		
Ala Leu Glu Pro Gly Gly Tyr Cys Gly Tyr Gly Xaa Cys Leu Trp Gly		
1	5	10 15
Glu Gly Thr Glu Val Thr Val Ser Ser		
	20	25
( 2 ) INFORMATION FOR SEQ ID NO:14:		
( i ) SEQUENCE CHARACTERISTICS:		
( A ) LENGTH: 23 amino acids		
( B ) TYPE: amino acid		
( C ) STRANDEDNESS: single		
( D ) TOPOLOGY: linear		
( i i ) MOLECULE TYPE: peptide		
( i x ) FEATURE:		
( A ) NAME/KEY: Domain		
( B ) LOCATION: 1..12		
( D ) OTHER INFORMATION: /label=CDR3		
( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:14:		
Val Ser Leu Met Asp Arg Ile Ser Gln His Gly Cys Arg Gly Gln Gly		
1	5	10 15
Thr Glu Val Thr Val Ser Leu		
	20	
( 2 ) INFORMATION FOR SEQ ID NO:15:		
( i ) SEQUENCE CHARACTERISTICS:		
( A ) LENGTH: 29 amino acids		
( B ) TYPE: amino acid		
( C ) STRANDEDNESS: single		
( D ) TOPOLOGY: linear		

-continued

---

```

(i i ) MOLECULE TYPE: peptide

(i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: L:18
      ( D ) OTHER INFORMATION: /label-CDR3

(x i ) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Val  Pro  Ala  His  Leu  Gly  Pro  Gly  Ala  Ile  Leu  Asp  Leu  Lys  Lys  Tyr
1          5          10          15

Lys  Tyr  Trp  Gly  Gln  Gly  Thr  Gln  Val  Thr  Val  Ser  Ser
          20          25

( 2 ) INFORMATION FOR SEQ ID NO:16:

(i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 26 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

(i i ) MOLECULE TYPE: peptide

(i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: L:15
      ( D ) OTHER INFORMATION: /label-CDR3

(x i ) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Phe  Cys  Tyr  Ser  Thr  Ala  Gly  Asp  Gly  Gln  Ser  Gly  Gln  Met  Tyr  Trp
1          5          10          15

Gly  Gln  Gly  Thr  Gln  Val  Thr  Val  Ser  Ser
          20          25

( 2 ) INFORMATION FOR SEQ ID NO:17:

(i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 26 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

(i i ) MOLECULE TYPE: peptide

(i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: L:15
      ( D ) OTHER INFORMATION: /label-CDR3

(x i ) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Gln  Leu  Ser  Gly  Gly  Ser  Cys  Gln  Leu  Pro  Leu  Leu  Phe  Asp  Tyr  Trp
1          5          10          15

Gly  Gln  Gly  Thr  Gln  Val  Thr  Val  Ser  Ser
          20          25

( 2 ) INFORMATION FOR SEQ ID NO:18:

(i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 28 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

(i i ) MOLECULE TYPE: peptide

(i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: L:17
      ( D ) OTHER INFORMATION: /label-CDR3

(x i ) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

-continued

```

A s p   T r p   L y s   T y r   T r p   T h r   C y s   G l y   A l a   G l n   T h r   G l y   G l y   T y r   P h e   G l y
1           5           10
G l n   T r p   G l y   G l n   G l y   A l a   G l n   V a l   T h r   V a l   S e r   S e r
                20           25

```

( 2 ) INFORMATION FOR SEQ ID NO:19:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 35 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: 1..24
- ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

A r g   L e u   T h r   G l n   M e t   G l y   A l a   C y s   A s p   A l a   A r g   T r p   A l a   T h r   L e u   A l a
1           5           10           15
T h r   A r g   T h r   P h e   A l a   T y r   A s n   T y r   T r p   G l y   G l n   G l y   T h r   G l n   V a l   T h r
                20           25           30
V a l   S e r   S e r
                35

```

( 2 ) INFORMATION FOR SEQ ID NO:20:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 27 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: 1..16
- ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

G l n   L y s   L y s   A s p   A r g   T h r   A r g   T r p   A l a   G l n   P r o   A r g   G l u   T r p   A s n   A s n
1           5           10           15
T r p   G l y   G l n   G l y   T h r   G l n   V a l   T h r   A l a   S e r   S e r
                20           25

```

( 2 ) INFORMATION FOR SEQ ID NO:21:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 32 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: 1..21
- ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```

G l y   S e r   A r g   P h e   S e r   S e r   P r o   V a l   G l y   S e r   T h r   S e r   A r g   L e u   G l u   S e r
1           5           10           15
S e r   A s p   T y r   A s n   T y r   T r p   G l y   G l n   G l y   T h r   G l n   V a l   T h r   A l a   S e r   S e r

```

-continued

20	25	30
( 2 ) INFORMATION FOR SEQ ID NO:22:		
( i ) SEQUENCE CHARACTERISTICS:		
( A ) LENGTH: 27 amino acids		
( B ) TYPE: amino acid		
( C ) STRANDEDNESS: single		
( D ) TOPOLOGY: linear		
( ii ) MOLECULE TYPE: peptide		
( ix ) FEATURE:		
( A ) NAME/KEY: Domain		
( B ) LOCATION: 1..16		
( D ) OTHER INFORMATION: /label=CDR3		
( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:22:		
Ala Asp Pro Ser Ile Tyr Tyr Ser Ile Leu Xaa Ile Glu Tyr Lys Tyr		
1 5 10 15		
Trp Gly Glu Gly Thr Glu Val Thr Val Ser Ser		
20 25		
( 2 ) INFORMATION FOR SEQ ID NO:23:		
( i ) SEQUENCE CHARACTERISTICS:		
( A ) LENGTH: 33 amino acids		
( B ) TYPE: amino acid		
( C ) STRANDEDNESS: single		
( D ) TOPOLOGY: linear		
( ii ) MOLECULE TYPE: peptide		
( ix ) FEATURE:		
( A ) NAME/KEY: Domain		
( B ) LOCATION: 1..22		
( D ) OTHER INFORMATION: /label=CDR3		
( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:23:		
Asp Ser Pro Cys Tyr Met Pro Thr Met Pro Ala Pro Pro Ile Arg Asp		
1 5 10 15		
Ser Phe Gly Trp Asp Asp Phe Gly Glu Gly Thr Glu Val Thr Val Ser		
20 25 30		
Ser		
( 2 ) INFORMATION FOR SEQ ID NO:24:		
( i ) SEQUENCE CHARACTERISTICS:		
( A ) LENGTH: 26 amino acids		
( B ) TYPE: amino acid		
( C ) STRANDEDNESS: single		
( D ) TOPOLOGY: linear		
( ii ) MOLECULE TYPE: peptide		
( ix ) FEATURE:		
( A ) NAME/KEY: Domain		
( B ) LOCATION: 1..15		
( D ) OTHER INFORMATION: /label=CDR3		
( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:24:		
Thr Ser Ser Phe Tyr Trp Tyr Cys Thr Thr Ala Pro Tyr Asa Val Trp		
1 5 10 15		
Gly Glu Gly Thr Glu Val Thr Val Ser Ser		
20 25		
( 2 ) INFORMATION FOR SEQ ID NO:25:		
( i ) SEQUENCE CHARACTERISTICS:		

-continued

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(i) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Domain
- (B) LOCATION: 1..16
- (D) OTHER INFORMATION: /label=CDR3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```

Thr Glu Ile Glu Trp Tyr Gly Cys Asn Leu Arg Thr Thr Phe Thr Arg
1           5           10           15
Trp Gly Glu Gly Thr Gln Val Thr Val Ser Ser
          20           25

```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Domain
- (B) LOCATION: 1..22
- (D) OTHER INFORMATION: /label=CDR3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```

Asn Gln Leu Ala Gly Gly Trp Tyr Leu Asp Pro Asn Tyr Trp Leu Ser
1           5           10           15
Val Gly Ala Tyr Ala Ile Trp Gly Gln Gly Thr His Val Thr Val Ser
          20           25           30
Ser

```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Domain
- (B) LOCATION: 1..24
- (D) OTHER INFORMATION: /label=CDR3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```

Arg Leu Thr Gln Met Gly Ala Cys Asp Ala Arg Trp Ala Thr Leu Ala
1           5           10           15
Thr Arg Thr Phe Ala Tyr Asn Tyr Trp Gly Arg Gly Thr Gln Val Thr
          20           25           30
Val Ser Ser
          35

```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

-continued

---

```

( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
  ( A ) NAME/KEY: Domain
  ( B ) LOCATION: 1-24
  ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:28:

  Asp Gly Trp Thr Arg Lys Glu Gly Gly Ile Gly Leu Pro Trp Ser Val
  1          5          10          15

  Glu Cys Glu Asp Gly Tyr Asa Tyr Trp Gly Glu Gly Thr Glu Val Thr
  20          25          30

  Val Ser Ser
  35

( 2 ) INFORMATION FOR SEQ ID NO:29:

  ( i ) SEQUENCE CHARACTERISTICS:
    ( A ) LENGTH: 21 amino acids
    ( B ) TYPE: amino acid
    ( C ) STRANDEDNESS: single
    ( D ) TOPOLOGY: linear

  ( i i ) MOLECULE TYPE: peptide

  ( i x ) FEATURE:
    ( A ) NAME/KEY: Domain
    ( B ) LOCATION: 1-10
    ( D ) OTHER INFORMATION: /label=CDR3

  ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:29:

  Asp Ser Tyr Pro Cys His Leu Leu Asp Val Trp Gly Glu Gly Thr Glu
  1          5          10          15

  Val Thr Val Ser Ser
  20

( 2 ) INFORMATION FOR SEQ ID NO:30:

  ( i ) SEQUENCE CHARACTERISTICS:
    ( A ) LENGTH: 23 amino acids
    ( B ) TYPE: amino acid
    ( C ) STRANDEDNESS: single
    ( D ) TOPOLOGY: linear

  ( i i ) MOLECULE TYPE: peptide

  ( i x ) FEATURE:
    ( A ) NAME/KEY: Domain
    ( B ) LOCATION: 1-12
    ( D ) OTHER INFORMATION: /label=CDR3

  ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:30:

  Val Glu Tyr Pro Ile Ala Asp Met Cys Ser Arg Tyr Gly Asp Pro Gly
  1          5          10          15

  Thr Glu Val Thr Val Ser Ser
  20

( 2 ) INFORMATION FOR SEQ ID NO:31:

  ( i ) SEQUENCE CHARACTERISTICS:
    ( A ) LENGTH: 27 amino acids
    ( B ) TYPE: amino acid
    ( C ) STRANDEDNESS: single
    ( D ) TOPOLOGY: linear

  ( i i ) MOLECULE TYPE: peptide

  ( i x ) FEATURE:

```

-continued

(A) NAME/KEY: Domain  
 (B) LOCATION: 1..27  
 (D) OTHER INFORMATION: /label=CH2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```

Ala  Pro  Glu  Leu  Leu  Gly  Gly  Pro  Thr  Val  Phe  Ile  Phe  Pro  Pro  Lys
 1              5              10
Pro  Lys  Asp  Val  Leu  Ser  Ile  Thr  Leu  Thr  Pro
              20              25

```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
 (A) NAME/KEY: Domain  
 (B) LOCATION: 1..27  
 (D) OTHER INFORMATION: /label=CH2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```

Ala  Pro  Glu  Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Thr  Lys
 1              5              10              15
Pro  Lys  Asp  Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro
              20              25

```

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
 (A) NAME/KEY: Domain  
 (B) LOCATION: 1..27  
 (D) OTHER INFORMATION: /label=CH2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```

Ala  Pro  Glu  Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Pro  Lys
 1              5              10              15
Pro  Lys  Asp  Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro
              20              25

```

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
 (A) NAME/KEY: Domain  
 (B) LOCATION: 1..27  
 (D) OTHER INFORMATION: /label=CH2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```

Ala  Pro  Glu  Leu  Leu  Gly  Gly  Pro  Ser  Val  Phe  Ile  Phe  Pro  Pro  Lys
 1              5              10              15

```



-continued

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro  
20 25

( 2 ) INFORMATION FOR SEQ ID NO:35:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 12 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: peptide

( ix ) FEATURE:  
 ( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1-12  
 ( D ) OTHER INFORMATION: /label=CD3

( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Gly Glu Thr Arg Glu Pro Gln Val Tyr Thr Leu Ala  
1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:36:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 18 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: peptide

( ix ) FEATURE:  
 ( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1-18  
 ( D ) OTHER INFORMATION: /label=CD3

( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Gly Gln Thr Arg Glu Pro Gln Val Tyr Thr Leu Ala Pro Xaa Arg Leu  
1 5 10 15

Glu Leu

( 2 ) INFORMATION FOR SEQ ID NO:37:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 12 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: peptide

( ix ) FEATURE:  
 ( A ) NAME/KEY: Region  
 ( B ) LOCATION: 1-12  
 ( D ) OTHER INFORMATION: /label=tlmgo

( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Gly Thr Asn Glu Val Cys Lys Cys Pro Lys Cys Pro  
1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:38:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 35 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: peptide

-continued

---

```

( i x ) FEATURE:
  ( A ) NAME/KEY: Region
  ( B ) LOCATION: 1-35
  ( D ) OTHER INFORMATION: /label=hiage

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:38:

  Glu  Pro  Lys  Ile  Pro  Gln  Pro  Gln  Pro  Lys  Pro  Gln  Pro  Gln  Pro  Gln
   1          5          10          15

  Pro  Gln  Pro  Lys  Pro  Gln  Pro  Lys  Pro  Gln  Pro  Gln  Cys  Thr  Cys  Pro
   20          25          30

  Lys  Cys  Pro
   35

( 2 ) INFORMATION FOR SEQ ID NO:39:

  ( i ) SEQUENCE CHARACTERISTICS:
    ( A ) LENGTH: 28 amino acids
    ( B ) TYPE: amino acid
    ( C ) STRANDEDNESS: single
    ( D ) TOPOLOGY: linear

  ( i i ) MOLECULE TYPE: peptide

  ( i x ) FEATURE:
    ( A ) NAME/KEY: Domain
    ( B ) LOCATION: 1-28
    ( D ) OTHER INFORMATION: /label=CH2

  ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:39:

  Ala  Pro  Gln  Leu  Leu  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Pro  Lys
   1          5          10          15

  Pro  Lys  Asp  Val  Leu  Ser  Ile  Ser  Gly  Xaa  Pro  Lys
   20          25

( 2 ) INFORMATION FOR SEQ ID NO:40:

  ( i ) SEQUENCE CHARACTERISTICS:
    ( A ) LENGTH: 28 amino acids
    ( B ) TYPE: amino acid
    ( C ) STRANDEDNESS: single
    ( D ) TOPOLOGY: linear

  ( i i ) MOLECULE TYPE: peptide

  ( i x ) FEATURE:
    ( A ) NAME/KEY: Domain
    ( B ) LOCATION: 1-28
    ( D ) OTHER INFORMATION: /label=CH2

  ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:40:

  Ala  Pro  Gln  Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Thr  Lys
   1          5          10          15

  Pro  Lys  Asp  Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro  Lys
   20          25

( 2 ) INFORMATION FOR SEQ ID NO:41:

  ( i ) SEQUENCE CHARACTERISTICS:
    ( A ) LENGTH: 28 amino acids
    ( B ) TYPE: amino acid
    ( C ) STRANDEDNESS: single
    ( D ) TOPOLOGY: linear

  ( i i ) MOLECULE TYPE: peptide

  ( i x ) FEATURE:
    ( A ) NAME/KEY: Domain
    ( B ) LOCATION: 1-28
    ( D ) OTHER INFORMATION: /label=CH2

```

-continued

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```

Ala Pro Glu Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro Pro Lys
1           5           10           15

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro Lys
                20           25

```

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Domain
- (B) LOCATION: L28
- (D) OTHER INFORMATION: /label=C12

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro Pro Lys
1           5           10           15

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro Lys
                20           25

```

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```

Val Thr Val Ser Ser Gly Thr Asa Glu Val Cys Lys Cys Pro Lys Cys
1           5           10           15

Pro Ala Pro Glu Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro
                20           25           30

```

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION: L54
- (D) OTHER INFORMATION: /label=tripe

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```

Val Thr Val Ser Ser Glu Pro Lys Ile Pro Glu Pro Glu Pro Lys Pro
1           5           10           15

Glu Pro Glu Pro Glu Pro Glu Pro Lys Pro Glu Pro Lys Pro Glu Pro
                20           25           30

Glu Cys Thr Cys Pro Lys Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro
                35           40           45

Ser Val Phe Ile Phe Pro
                50

```

-continued

( 2 ) INFORMATION FOR SEQ ID NO:45:

- ( i ) SEQUENCE CHARACTERISTICS:
  - ( A ) LENGTH: 14 amino acids
  - ( B ) TYPE: amino acid
  - ( C ) STRANDEDNESS: single
  - ( D ) TOPOLOGY: linear
- ( i i ) MOLECULE TYPE: peptide
- ( i x ) FEATURE:
  - ( A ) NAME/KEY: Region
  - ( B ) LOCATION: 1..14
  - ( D ) OTHER INFORMATION: /label=kingo
- ( i x ) FEATURE:
  - ( A ) NAME/KEY: Domain
  - ( B ) LOCATION: 1..14
  - ( D ) OTHER INFORMATION: /label=CH2
- ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```

Ala Pro Glu Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro
1           5           10

```

( 2 ) INFORMATION FOR SEQ ID NO:46:

- ( i ) SEQUENCE CHARACTERISTICS:
  - ( A ) LENGTH: 14 amino acids
  - ( B ) TYPE: amino acid
  - ( C ) STRANDEDNESS: single
  - ( D ) TOPOLOGY: linear
- ( i i ) MOLECULE TYPE: peptide
- ( i x ) FEATURE:
  - ( A ) NAME/KEY: Domain
  - ( B ) LOCATION: 1..14
  - ( D ) OTHER INFORMATION: /label=CH2
- ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro
1           5           10

```

( 2 ) INFORMATION FOR SEQ ID NO:47:

- ( i ) SEQUENCE CHARACTERISTICS:
  - ( A ) LENGTH: 21 base pairs
  - ( B ) TYPE: nucleic acid
  - ( C ) STRANDEDNESS: single
  - ( D ) TOPOLOGY: linear
- ( i i ) MOLECULE TYPE: Other
  - ( A ) DESCRIPTION: DNA (synthetic)
- ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CGCCATCAAG GTAACA GTTG A

21

( 2 ) INFORMATION FOR SEQ ID NO:48:

- ( i ) SEQUENCE CHARACTERISTICS:
  - ( A ) LENGTH: 22 base pairs
  - ( B ) TYPE: nucleic acid
  - ( C ) STRANDEDNESS: single
  - ( D ) TOPOLOGY: linear
- ( i i ) MOLECULE TYPE: Other
  - ( A ) DESCRIPTION: DNA (synthetic)
- ( i x ) FEATURE:
  - ( A ) NAME/KEY: min\_feature
  - ( B ) LOCATION: 12..17
  - ( D ) OTHER INFORMATION: /label=33rd site

-continued

---

```

/ note= "RESTRICTION SITE"

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:48:
A G G T C C A G C T   G C T C G A G T C T   G G                               2 2

( 2 ) INFORMATION FOR SEQ ID NO:49:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 22 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( i x ) FEATURE:
      ( A ) NAME/KEY: misc_feature
      ( B ) LOCATION: 12..17
      ( D ) OTHER INFORMATION: fabeli-OHcol site
                               / note= "Restriction site"

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:49:
A G C T C C A G C T   G C T C G A G T C T   G G                               2 2

( 2 ) INFORMATION FOR SEQ ID NO:50:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 22 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( i x ) FEATURE:
      ( A ) NAME/KEY: misc_feature
      ( B ) LOCATION: 12..17
      ( D ) OTHER INFORMATION: fabeli-OHcol site
                               / note= "Restriction site"

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:50:
A G G T C C A G C T   T C T C G A G T C T   G G                               2 2

( 2 ) INFORMATION FOR SEQ ID NO:51:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 28 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:51:
T C T T A A C T A G   T G A G G A G A C G   G T G A C C T G                       2 8

( 2 ) INFORMATION FOR SEQ ID NO:52:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 30 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( i x ) FEATURE:
      ( A ) NAME/KEY: misc_feature

```

-continued

---

```

      ( B ) LOCATION: 1.5
      ( D ) OTHER INFORMATION: /label=/Spel

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:52:
CTAGTGCAACC ACCATCACCA TCACTAATAAG
3 0

( 2 ) INFORMATION FOR SEQ ID NO:53:

  ( i ) SEQUENCE CHARACTERISTICS:
    ( A ) LENGTH: 30 base pairs
    ( B ) TYPE: nucleic acid
    ( C ) STRANDEDNESS: single
    ( D ) TOPOLOGY: linear

  ( i i ) MOLECULE TYPE: Other
    ( A ) DESCRIPTION: DNA (synthetic)

  ( i x ) FEATURE:
    ( A ) NAME/KEY: misc_feature
    ( B ) LOCATION: 1..30
    ( D ) OTHER INFORMATION: /note= "Sequence complementary to
      SEQ ID NO: 52"

  ( i x ) FEATURE:
    ( A ) NAME/KEY: misc_feature
    ( B ) LOCATION: 26..30
    ( D ) OTHER INFORMATION: /label=/ExoR

  ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:53:
ACGTGGTGGT AGTGGTAOTG ATTATCTTAA
3 0

( 2 ) INFORMATION FOR SEQ ID NO:54:

  ( i ) SEQUENCE CHARACTERISTICS:
    ( A ) LENGTH: 43 amino acids
    ( B ) TYPE: amino acid
    ( C ) STRANDEDNESS: single
    ( D ) TOPOLOGY: linear

  ( i i ) MOLECULE TYPE: peptide

  ( v ) FRAGMENT TYPE: N-terminal

  ( v i ) ORIGINAL SOURCE:
    ( A ) ORGANISM: Cannabis dracunculoides

  ( i x ) FEATURE:
    ( A ) NAME/KEY: Domain
    ( B ) LOCATION: 1..25
    ( D ) OTHER INFORMATION: /label=/CH2

  ( i x ) FEATURE:
    ( A ) NAME/KEY: Domain
    ( B ) LOCATION: 26..43
    ( D ) OTHER INFORMATION: /label=/CH3

  ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:54:
Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Pro  Lys  Pro  Lys  Asp
1          5          10          15

Val  Leu  Ser  Ile  Xaa  Gly  Xaa  Pro  Lys  Gly  Gln  Thr  Arg  Gln  Pro  Gln
20          25          30

Val  Tyr  Thr  Leu  Ala  Pro  Xaa  Arg  Leu  Gln  Leu
35          40

```

( 2 ) INFORMATION FOR SEQ ID NO:55:

```

  ( i ) SEQUENCE CHARACTERISTICS:
    ( A ) LENGTH: 24 amino acids
    ( B ) TYPE: amino acid
    ( C ) STRANDEDNESS: single
    ( D ) TOPOLOGY: linear

```

-continued

---

```

(i i ) MOLECULE TYPE: peptide
(v ) FRAGMENT TYPE: N-terminal

(v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Caninus dromedarius

(i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1-24
      ( D ) OTHER INFORMATION: /label=CH2
              / note="Close #72/1"

(x i ) SEQUENCE DESCRIPTION: SEQ ID NO:55:

  Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Thr  Lys  Pro  Lys  Asp
   1              5              10              15

  Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro
                20

( 2 ) INFORMATION FOR SEQ ID NO:56:

(i i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 24 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

(i i ) MOLECULE TYPE: peptide
(v ) FRAGMENT TYPE: N-terminal

(v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Caninus dromedarius

(i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1-24
      ( D ) OTHER INFORMATION: /label=CH2

(x i ) SEQUENCE DESCRIPTION: SEQ ID NO:56:

  Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Pro  Lys  Pro  Lys  Asp
   1              5              10              15

  Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro
                20

( 2 ) INFORMATION FOR SEQ ID NO:57:

(i i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 24 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

(i i ) MOLECULE TYPE: peptide
(v ) FRAGMENT TYPE: N-terminal

(v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Caninus dromedarius

(i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1-24
      ( D ) OTHER INFORMATION: /label=CH2

(x i ) SEQUENCE DESCRIPTION: SEQ ID NO:57:

  Leu  Leu  Gly  Gly  Pro  Ser  Val  Phe  Ile  Phe  Pro  Pro  Lys  Pro  Lys  Asp
   1              5              10              15

  Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro
                20

```

-continued

## ( 2 ) INFORMATION FOR SEQ ID NO:58:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 30 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( v ) FRAGMENT TYPE: N-terminal

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Camelus dromedarius

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1-30
      ( D ) OTHER INFORMATION: /label=Framework 1
           / note="CAMEL"

```

## ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```

A s p   V a l   G l u   L e u   V a l   A l a   S e r   G l y   G l y   G l y   S e r   V a l   G l y   A l a   G l y   G l y
1               5               10               15
S e r   L e u   A r g   L e u   S e r   C y s   T h r   A l a   S e r   G l y   A s p   S e r   P h e   S e r
                20               25               30

```

## ( 2 ) INFORMATION FOR SEQ ID NO:59:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Camelus dromedarius

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1-11
      ( D ) OTHER INFORMATION: /label=Framework 4

```

## ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```

T r p   G l y   A r g   G l y   T h r   G l u   V a l   T h r   V a l   S e r   S e r
1               5               10

```

## ( 2 ) INFORMATION FOR SEQ ID NO:60:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Camelus dromedarius

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1-11
      ( D ) OTHER INFORMATION: /label=Framework 4

```

## ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```

T r p   G l y   G l u   G l y   T h r   H i s   V a l   T h r   V a l   S e r   S e r
1               5               10

```

## ( 2 ) INFORMATION FOR SEQ ID NO:61:



-continued

---

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..11
      ( D ) OTHER INFORMATION: /label=Framework 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:61:
  Trp  Gly  Gln  Gly  Ile  Gln  Val  Thr  Ala  Ser  Ser
   1          5          10

( 2 ) INFORMATION FOR SEQ ID NO:62:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 14 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1..14
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..14
      ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:62:
  Ala  Leu  Gln  Pro  Gly  Gly  Tyr  Cys  Gly  Tyr  Gly  Xaa  Cys  Leu
   1          5          10

( 2 ) INFORMATION FOR SEQ ID NO:63:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 12 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Canis lupus familiaris

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1..12
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..12
      ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:63:
  Val  Ser  Leu  Met  Asp  Arg  Ile  Ser  Gln  His  Gly  Cys
   1          5          10

( 2 ) INFORMATION FOR SEQ ID NO:64:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 18 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single

```

-continued

---

```

      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Camelus dromedarius

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1..18
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..18
      ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Val  Pro  Ala  His  Leu  Gly  Pro  Gly  Ala  Ile  Leu  Asp  Leu  Lys  Lys  Tyr
 1          5          10          15

Lys  Tyr

( 2 ) INFORMATION FOR SEQ ID NO:65:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 15 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Camelus bactrianus

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1..15
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..15
      ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Phe  Cys  Tyr  Ser  Thr  Ala  Gly  Asp  Gly  Gly  Ser  Gly  Glu  Met  Tyr
 1          5          10          15

( 2 ) INFORMATION FOR SEQ ID NO:66:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 15 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Camelus dromedarius

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1..15
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..15
      ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Glu  Leu  Ser  Gly  Gly  Ser  Cys  Glu  Leu  Pro  Leu  Leu  Phe  Asp  Tyr

```

-continued

1	5	10	15
( 2 ) INFORMATION FOR SEQ ID NO:67:			
( i ) SEQUENCE CHARACTERISTICS:			
( A ) LENGTH: 17 amino acids			
( B ) TYPE: amino acid			
( C ) STRANDEDNESS: single			
( D ) TOPOLOGY: linear			
( i i ) MOLECULE TYPE: protein			
( i x ) FEATURE:			
( A ) NAME/KEY: Region			
( B ) LOCATION: 1..17			
( D ) OTHER INFORMATION: /label=VH			
( i x ) FEATURE:			
( A ) NAME/KEY: Domain			
( B ) LOCATION: 1..17			
( D ) OTHER INFORMATION: /label=CDR3			
( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:67:			
A s p T r p L y s T y r T r p T h r C y s G l y A l a G l u T h r G l y G l y T y r P h e G l y			
1 5 10 15			
G l u			
( 2 ) INFORMATION FOR SEQ ID NO:68:			
( i ) SEQUENCE CHARACTERISTICS:			
( A ) LENGTH: 24 amino acids			
( B ) TYPE: amino acid			
( C ) STRANDEDNESS: single			
( D ) TOPOLOGY: linear			
( i i ) MOLECULE TYPE: protein			
( i x ) FEATURE:			
( A ) NAME/KEY: Region			
( B ) LOCATION: 1..24			
( D ) OTHER INFORMATION: /label=VH			
( i x ) FEATURE:			
( A ) NAME/KEY: Domain			
( B ) LOCATION: 1..24			
( D ) OTHER INFORMATION: /label=CDR3			
( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:68:			
A r g L e u T h r G l u M e t G l y A l a C y s A s p A l a A r g T r p A l a T h r L e u A l a			
1 5 10 15			
T h r A r g T h r P h e A l a T y r A s n T y r			
20			
( 2 ) INFORMATION FOR SEQ ID NO:69:			
( i ) SEQUENCE CHARACTERISTICS:			
( A ) LENGTH: 16 amino acids			
( B ) TYPE: amino acid			
( C ) STRANDEDNESS: single			
( D ) TOPOLOGY: linear			
( i i ) MOLECULE TYPE: protein			
( i x ) FEATURE:			
( A ) NAME/KEY: Region			
( B ) LOCATION: 1..16			
( D ) OTHER INFORMATION: /label=VH			
( i x ) FEATURE:			
( A ) NAME/KEY: Domain			
( B ) LOCATION: 1..16			
( D ) OTHER INFORMATION: /label=CDR3			

-continued

( 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:69:

Gln	Lys	Lys	Asp	Arg	Thr	Arg	Trp	Ala	Gln	Pro	Arg	Glu	Trp	Asn	Asn
1			5						10					15	

( 2 ) INFORMATION FOR SEQ ID NO:70:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 21 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:

- ( A ) NAME/KEY: Region
- ( B ) LOCATION: L:21
- ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: L:21
- ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Gly	Ser	Arg	Phe	Ser	Ser	Pro	Val	Gly	Ser	Thr	Ser	Arg	Leu	Glu	Ser
1				5				10						15	
Ser Asp Tyr Asn Tyr															
20															

( 2 ) INFORMATION FOR SEQ ID NO:71:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 16 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:

- ( A ) NAME/KEY: Region
- ( B ) LOCATION: L:16
- ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: L:16
- ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:71:

Ala	Asp	Pro	Ser	Ile	Tyr	Tyr	Ser	Ile	Leu	Xaa	Ile	Glu	Tyr	Lys	Tyr
1				5					10					15	

( 2 ) INFORMATION FOR SEQ ID NO:72:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 22 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:

- ( A ) NAME/KEY: Region
- ( B ) LOCATION: L:22
- ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: L:22
- ( D ) OTHER INFORMATION: /label=CDR3

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:72:

```

Asp Ser Pro Cys Tyr Met Pro Thr Met Pro Ala Pro Pro Ile Arg Asp
1           5           10           15

Ser Phe Gly Trp Asp Asp
20

```

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION: 1-15
- (D) OTHER INFORMATION: /label=VH

(ix) FEATURE:

- (A) NAME/KEY: Domain
- (B) LOCATION: 1-15
- (D) OTHER INFORMATION: /label=CDR3

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:73:

```

Thr Ser Ser Phe Tyr Trp Tyr Cys Thr Thr Ala Pro Tyr Asn Val
1           5           10           15

```

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION: 1-16
- (D) OTHER INFORMATION: /label=VH

(ix) FEATURE:

- (A) NAME/KEY: Domain
- (B) LOCATION: 1-16
- (D) OTHER INFORMATION: /label=CDR3

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:74:

```

Thr Glu Ile Glu Trp Tyr Gly Cys Asn Leu Arg Thr Thr Phe Thr Arg
1           5           10           15

```

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION: 1-22
- (D) OTHER INFORMATION: /label=VH

(ix) FEATURE:

- (A) NAME/KEY: Domain
- (B) LOCATION: 1-22

-continued

(D) OTHER INFORMATION: /label=CDR3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

```

A s n   G l a   L e u   A l a   G l y   G l y   T r p   T y r   L e u   A s p   P r o   A s a   T y r   T r p   L e u   S e r
 1             5             10             15

V a l   G l y   A l a   T y r   A l a   I l e
                20

```

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION: 1-24
- (D) OTHER INFORMATION: /label=VH

(ix) FEATURE:

- (A) NAME/KEY: Domain
- (B) LOCATION: 1-24
- (D) OTHER INFORMATION: /label=CDR3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

```

A r g   L e u   T h r   G l u   M e t   G l y   A l a   C y s   A s p   A l a   A r g   T r p   A l a   T h r   L e u   A l a
 1             5             10             15

T h r   A r g   T h r   P h e   A l a   T y r   A s a   T y r
                20

```

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION: 1-24
- (D) OTHER INFORMATION: /label=VH

(ix) FEATURE:

- (A) NAME/KEY: Domain
- (B) LOCATION: 1-24
- (D) OTHER INFORMATION: /label=CDR3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

```

A s p   G l y   T r p   T h r   A r g   L y s   G l u   G l y   G l y   I l e   G l y   L e u   P r o   T r p   S e r   V a l
 1             5             10             15

G l a   C y s   G l u   A s p   G l y   T y r   A s a   T y r
                20

```

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

-continued

---

```

      (A) NAME/KEY: Region
      (B) LOCATION: L:10
      (D) OTHER INFORMATION: /label=VH

  (ix) FEATURE:
      (A) NAME/KEY: Domain
      (B) LOCATION: L:10
      (D) OTHER INFORMATION: /label=CDR3

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:
      Asp Ser Tyr Pro Cys His Leu Leu Asp Val
      1           5           10

  (2) INFORMATION FOR SEQ ID NO:79:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 12 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein

      (ix) FEATURE:
          (A) NAME/KEY: Region
          (B) LOCATION: L:12
          (D) OTHER INFORMATION: /label=VH

      (ix) FEATURE:
          (A) NAME/KEY: Domain
          (B) LOCATION: L:12
          (D) OTHER INFORMATION: /label=CDR3

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:
      Val Glu Tyr Pro Ile Ala Asp Met Cys Ser Arg Tyr
      1           5           10

  (2) INFORMATION FOR SEQ ID NO:80:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 26 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein

      (vi) ORIGINAL SOURCE:
          (A) ORGANISM: Canis lupus familiaris

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:
      Glu Val Glu Leu Val Glu Ser Gly Gly Gly Leu Val Glu Pro Gly Gly
      1           5           10           15
      Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly
      20           25

  (2) INFORMATION FOR SEQ ID NO:81:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 14 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein

      (vi) ORIGINAL SOURCE:
          (A) ORGANISM: Canis lupus familiaris

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:
      Trp Val Arg Glu Ala Pro Gly Lys Gly Leu Glu Trp Val Ser
      1           5           10

```

## ( 2 ) INFORMATION FOR SEQ ID NO:82:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 32 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

## ( ii ) MOLECULE TYPE: protein

## ( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:82:

```

Arg  Phe  Thr  Ile  Ser  Arg  Asp  Asa  Ser  Lys  Asa  Thr  Leu  Tyr  Leu  Glu
1          5          10          15
Met  Asa  Ser  Leu  Arg  Ala  Glu  Asp  Thr  Ala  Val  Tyr  Tyr  Cys  Ala  Arg
          20          25          30

```

## ( 2 ) INFORMATION FOR SEQ ID NO:83:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 37 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

## ( ii ) MOLECULE TYPE: protein

## ( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:83:

```

Trp  Gly  Glu  Gly  Thr  Leu  Val  Thr  Val  Ser  Ser  Gly  Thr  Asa  Glu  Val
1          5          10          15
Cys  Lys  Cys  Pro  Lys  Cys  Pro  Ala  Pro  Glu  Leu  Pro  Gly  Gly  Pro  Ser
          20          25          30
Val  Phe  Val  Phe  Pro
          35

```

## ( 2 ) INFORMATION FOR SEQ ID NO:84:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 18 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

## ( ii ) MOLECULE TYPE: protein

## ( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:84:

```

Gly  Gly  Ser  Val  Glu  Gly  Gly  Gly  Ser  Leu  Arg  Leu  Ser  Cys  Ala  Ile
1          5          10          15
Ser  Gly

```

## ( 2 ) INFORMATION FOR SEQ ID NO:85:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 14 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

## ( ii ) MOLECULE TYPE: protein

## ( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:85:

```

Trp  Phe  Arg  Glu  Gly  Pro  Gly  Lys  Glu  Arg  Glu  Gly  Ile  Ala
1          5          10

```

## ( 2 ) INFORMATION FOR SEQ ID NO:86:

## ( i ) SEQUENCE CHARACTERISTICS:



-continued

(A) LENGTH: 32 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(i) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

```

Arg Phe Thr Ile Ser Gln Asp Ser Thr Leu Lys Thr Met Tyr Leu Leu
1           5           10           15
Met Asn Asn Leu Lys Pro Glu Asp Thr Gly Thr Tyr Tyr Cys Ala Ala
                20           25           30

```

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(i) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

```

Trp Gly Glu Gly Thr Glu Val Thr Val Ser Ser Glu Pro Lys Ile Pro
1           5           10           15
Glu Pro Glu Pro Lys Pro Glu Pro Glu Pro Glu Pro Glu Pro Lys Pro
                20           25           30
Glu Pro Lys Pro Glu Pro Glu Cys Thr Cys Pro Lys Cys Pro Ala Pro
                35           40           45
Glu Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro
                50           55           60

```

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(i) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

```

Gly Gly Ser Val Glu Ala Glu Gly Ser Leu Arg Leu Ser Cys Ala Ser
1           5           10           15
Ser Ser

```

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(i) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

```

Trp Tyr Arg Glu Ala Pro Gly Lys Glu Arg Glu Phe Val Ser
1           5           10

```

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 amino acids  
 (B) TYPE: amino acid

-continued

( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:90:

Arg	Phe	Thr	Ile	Ser	Gln	Asp	Ser	Ala	Lys	Asn	Thr	Val	Tyr	Leu	Gln
1				5					10					15	
Met	Asa	Ser	Leu	Lys	Pro	Glu	Asp	Thr	Ala	Met	Tyr	Tyr	Cys	Lys	Ile
			20					25					30		

( 2 ) INFORMATION FOR SEQ ID NO:91:

( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 37 amino acids  
( B ) TYPE: amino acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:91:

Trp	Gly	Gln	Gly	Thr	Gln	Val	Thr	Val	Ser	Ser	Gly	Thr	Asa	Glu	Val
1				5				10					15		
Cys	Lys	Cys	Pro	Lys	Cys	Pro	Ala	Pro	Glu	Leu	Pro	Gly	Gly	Pro	Ser
			20				25					30			
Val	Phe	Val	Phe	Pro											
			35												

( 2 ) INFORMATION FOR SEQ ID NO:92:

( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 400 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:92:

CTCGAATCTG	GGGGAGGATC	GGTGCAAGCT	GGAGGGTCTC	TGAAGACTCTC	GTGCGCAGCC	60
TCTGGATACA	GTAATTGTCC	CCTCACTTGG	AGCTGGTATC	GCCAATTTC	AGGAACGGAG	120
CGCGAATTCG	TCTCCAATAT	GAATCCGGAAT	GGAAATACCA	AGTACACATA	CTCCGTGAAG	180
GGCCGCTTCA	CCATGTCCCG	AGGCAGCACC	GAATACACAG	TATTCTGCA	AATGGACAAT	240
CTGAAACCTG	AGGACACGGC	GATGTATTAC	TGTAAACACG	CCCTACACCC	TGGGGGTTAT	300
TGTGGGTATG	GGTANTGCTT	CTGGGGCCAG	GGGACCCAGG	TCACCGTCTC	CTCACTAGTT	360
ACCCGATACG	CGTTCCGGAC	TACGGTCTCT	AATAGAAATC			400

( 2 ) INFORMATION FOR SEQ ID NO:93:

( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 391 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:93:

CTCGAATCTG	GGGGAGGCTC	GGTGCAAGCT	GGAGGGTCTC	TGAAGACTCTC	CTGTGATCT	60
TCTTCTAAAT	ATATGCTTGG	CACCTACGAC	ATGACTGGT	ACCGCCAGGC	TCCAGGCCAG	120
GAAGCCGAAAT	TGTCTCTAAG	TATAAATATT	GATGGTAAAG	CAACATACGC	AGACTCCGTG	180

-continued

---

```

AAGGGCCGAT TCACCATCTC CCAAGACAGC GCCAAGAAC CCGTGTATCT GCAGATGAAC      240
AGCCTGAAAC CTGAGGACAC GCGATGTAT TACTGTAAAA TAGATTCTGA CCCGTGCCAT      300
CTCCTTGATG TCTGGGCCCA GGGGACCCAG GTCACCGTCT CTTACTACTG TACCCTGACG      360
AGCTTCGGA CTACGGTTCT TAATAGAATT C

```

( 2 ) INFORMATION FOR SEQ ID NO-94:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 443 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

```

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO-94:

```

CAGGTGAAC TGCTCGATC TGGAGGAGGC TCGGTGCA GA CTGGAAGATC TCTGAGACTC      60
TCTGTGCA G TCTCTGGATT CTCCTTTAGT ACCAGTTGTA TGCCCTGGTT CCGCCAGGCT      120
TCAGGAAAGC AGCGTAGGG GGTCCGACGC ATTAATAAGT GCGGTGGTAG GACATACTAC      180
AACACATATG TCGCCGATC CGTGAAGGGC GATTCGCCA TCTCCCAAGG CAACGCCAAG      240
ACCACGGTAT ATCTTGATAT GAACAACCTA ACCCTGAA G ACACGGCTAC GTATTACTGT      300
GCGCGGCTCC CAGCCCACTT GGGACCTGGC GCCATCTTG ATTTGAAAAA GTATAAGTAC      360
TGGGGCCAGG GGACCCAGGT CACCGTCTCC TCACTAGCTA GTTACCCGTA CGACGTTCCG      420
GACTACGGT CTTAATAGAA TTC

```

( 2 ) INFORMATION FOR SEQ ID NO-95:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 433 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

```

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO-95:

```

CTCGAATCTG GGGGAGGTC GGTGCAAGCT GGAAGGCTCT TGAGACTCTC CTGTAATGTC      60
TCTGGCTCTC CCAAGTAGTAC TTATTGCTG GGTGCTGTC GCCAAGCTCC AGGGAGGGAG      120
CGTGAGGGGG TCACAGCGAT TAACACTGAT GCGAGTATCA TATACGACG CAGACTCGTG      180
AAGGCCGAT TCACCATCTC CCAAGACACC GCCAAGGAAA CCGTACATCT CCAGATGAAC      240
AACCTGCAAC CTGAGGATAC GGGCACTTAT TACTGCGCG CAAAGACTGAC GGAAGTGGGG      300
GCTTGATG CAGAGTGGGC GACCTTAGCG ACAAGGACGT TTGCTATAA CTACTGGGGC      360
CGGGGGACCC AGGTCAACGT CTCCTCACTA GTTACCCGTA CGACGTTCCG GACTACGGTT      420
CTTAATAGAA TTC

```

( 2 ) INFORMATION FOR SEQ ID NO-96:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 440 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

```

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO-96:

-continued

CAGGTGAAAC	TGCTCGAGTC	TGGGGGAGGG	TCGGTGCAGG	CTGGAGGGTC	TCGAGACTC	60
TCCTGTAAAT	TCCTCTGGTC	TCCCAAGTAGT	ACTTATTGCC	TGGGCTGGTT	CCGCCAGGCT	120
CCAGGGAAAG	AGCGTGAGGG	GGTCACAGCG	ATTAACTACTG	ATGGCAGTGT	CATATACGCA	180
GCCGACTCCG	TGAAAGGGCCG	ATTACACCATC	TCCCAAAGACA	CCGCCAAGAA	AACGGTATAT	240
CTCCAGATGA	ACAACCTGCA	ACCTGAGGAT	ACGGCCACCT	ATTACTGCGC	GGCAAGACTG	300
ACGAGATAGG	GGGCTTGTGA	TGCGAGATGG	GCAGCCTTAG	CGACAAGGAC	GTTTGCCTAT	360
GTCTACTGGG	GCCGGGGGAC	CCAAGTCAAC	GTCTCTCTAC	TAGCTAGTGA	CCGCTACGAC	420
GTTCGGGACT	ACGGTTCCTTA	ATAGAATTG				449

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 424 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CTCGAGTCTG	GAGGAGGCTC	GGCCCAAGCT	GGAGGATCTC	TGAGACTCTC	CTGTGCAGCC	60
CACGGGATTC	CGCTCAATGG	TTACTACATC	GCCTGGTTCC	GTCAAGGCTCC	TGGGAAAGGGG	120
CGTGAGGGGG	TCGCAACAAT	TAATGGTGGT	CGCGACGTCA	CATACACTGC	CGACTCCGTG	180
ACGGGCCGAT	TTACCATCTC	CCGAGACAGC	CCCAAGGAATA	CGGTGTATCT	GCAAGATGAAC	240
AGCCTGAAAC	CTGAGGACAC	GGCCATCTAC	TTCTGTGCAG	CAGGCTCGCG	TTTTCTAGT	300
CCTGTGGGGA	GCACCTTCTAG	ACTCGAAAGT	AGCGACTATA	ACTATTGGGG	CCAGGGGATC	360
CAGGTCAACC	TCACCTCACT	AGTTACCCGT	ACGACGTTC	GGACTACGGT	TCITTAATAGA	420
ATTTC						424

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 413 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO:98:

CTCGAGTCTG	GAGGAGGCTC	GGTTCAGGCT	GGAGGGTCCC	TTAAGACTCTC	CTGTGCAGCC	60
TCGTACTACA	CCATCACTGA	TTATTGCATG	GCCTGGTTCC	GCCAAGGCTCC	AGGGAAAGGAG	120
CGTGAAATTGG	TCGACAGCGAT	TCAAGTTGTC	CGTAGTGATA	CTGCGCTCAC	AGACTACGCC	180
GACTCCGTGA	AGGGACGATT	CACCATCTCC	CAAAGGCAACA	CCAAGAACAC	AGTGAATCTG	240
CAAAATGAACA	GCCTGACACC	TGAAGACACG	GCCATCTACA	GTGTGTGGGC	AACCAAGTAGT	300
TTTTACTGGT	ACTGACCCAC	GGCGCCTTAT	AACGTCTGGG	GTCAAGGGGAC	CCAGGCTCAC	360
GTCTCTCTAC	TAGTTACCCG	TACGACGTTT	CGGACTACGG	TTCITTAATAG	AATTTC	413

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 406 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

-continued

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:99:

CTCGAGTCTG	GGGGAGGCTC	GGTGCAGGCT	GGAGGGTCTC	TGAGACTCTC	CTGTGCAATC	60
TCTGGATACA	CGTACGGTAG	CTTCTGTATG	GGCTGGTTCC	GCGAGGGTCC	AGGCAAGGAA	120
CGTGAGGGGA	TCGCAACTAT	TCITAATGGT	GGTACTAACA	CATACTATGC	CGACTCGGTG	180
AAGGGCCGAT	TCACCATCTC	CCAAGACAGC	ACGTTGAAAG	CGATGTATCT	GCTAATGAAC	240
AACCTGAAAC	CTGAAGACAC	GGGCACCTAT	TACTGTGCTG	CAGAAGTAA	TGGTGGTAGT	300
TGTGAATTGC	CTTTGCTATT	TGACTACTGG	GGCCAGGGCA	CCCAGGTAC	CGTCTCCTCA	360
CTAGTTACCC	GTACGACGTT	CCGGACTACG	GTTCCTTAATA	GAATTC		406

( 2 ) INFORMATION FOR SEQ ID NO:100:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 427 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:100:

CTCGAGTCTG	GGGGAGGCTC	GGTGCAGGCT	GGAGGGTCTC	TGAGACTCTC	CTGTACAGGC	60
TCTGGATACA	CCTATAGTAC	CTTCTGTCTG	GGGTGGTTCC	GCCAGGCTCC	AGGGAAGGAG	120
CGTGAGGGGG	TGCGGGGTAT	TAATAGTGCA	GGAGGTAATA	CTTACTATGC	CGACGCCGTG	180
AAGGGCCGAT	TCACCATCTC	CCAAGGGAAT	GCCAAAGATA	CGGTGTTTCT	GCAATGGAT	240
AACCTGAAAC	CTGAGGACAC	GGCCATCTAT	TACTGCGCGG	CGGATAGTCC	ATGTTACATG	300
CCGACTATGC	CCGCTCCCCC	GATACGAGAC	AAGTTTGCTC	GGGATGATTT	TGGCCAGGGG	360
ACCCAGGTCA	CCGTCTCCTC	ACTAGTTACC	CGTACGACGT	TCCGGACTAC	GGTTCTTAAT	420
AGAATTTC						427

( 2 ) INFORMATION FOR SEQ ID NO:101:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 409 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CTCGAGTCA	GGGGAGGCTC	GGTACAGGTT	GGAGGGTCTC	TGAGACTCTC	CTGTGTAGCC	60
TCTACTACAC	CCGACAGTAG	CACCTGTATA	GGCTGGTTCC	GCCAGGCTCC	AGGGAAGGAG	120
CGCGAGGGGG	TCGCAAGTAT	ATATTTTGGT	GATGGTGGTA	CGAATTATCG	CGACTCCGTG	180
AAGGGCCGAT	TCACCATCTC	CCAACCTAAC	GCCACAGAAC	CAGTGTATCT	GCAAAATGAAC	240
AGCCTGAAAC	CTGAGGACAG	CGCCATGTAC	TACTGTGCAA	TCACTGAAAT	TGAGTGGTAT	300
GGGTGCAATT	TAAGGACTAC	TTTACTCGC	TGGGGCCAGG	GGACCCAGGT	CACCGTCTCC	360
TCACTAGTTA	CCCGTACGAC	GTTCGGGACT	ACGGTTCTTA	ATAGAAATTC		409

( 2 ) INFORMATION FOR SEQ ID NO:102:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 445 base pairs

( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( x i ) SEQUENCE DESCRIPTION: SBQ ID NO:102:

TCTGAGTCTG	GGGGAGGCTC	GGTACAAACT	GGAGGGTCTC	TGAGACTCTC	TTGCGAAATC	60
TCTGGATTGA	CTTTTGATTG	TTCGATGATG	GGGTGGTACC	CCCAAGCTCC	AGGGGATGAG	120
TGCAAAATCG	TCTCAGGTAT	TCTGAGTCGT	GGTACTCCAT	ATACAAAAGG	TGGAGACTAT	180
TCTGAGTCTG	TGAGGGGCCG	GGTACCATC	TCCAGAGACA	ACGCCAAGAA	TGGATATATC	240
CTTCAAATGA	ACGACCTGAA	ACCTGAGGAC	ACGGCCATGT	ATTACTGCGC	GGTAGATGGT	300
TGGACCCGGA	AGGAAGGGGG	AATCGGGTTA	CCCTGGTCCG	TCCAAATGTA	AGATGGTTAT	360
AACTATTGGG	CCCAAGGGAC	CCAGGTACCC	GTCTCTCTAC	TAGTATACCC	GCACAGCTTC	420
CGGACTACGG	TTCTTAATAG	AATTC				445

( 2 ) INFORMATION FOR SEQ ID NO:103:

( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 394 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:103:

TCTGAGTCTG	GAGGAGGCTC	GGTGCAGGCT	GGAAGGTTCT	TGAGACTCTC	CTGTGTAGCC	60
TCTGGATTCA	ATTTGGAAC	TTCTCGTATG	CGGTGGTACC	GCCAGACTCC	AGGAAATGTG	120
TCTGAGTTGG	TCTCAAGTAT	TAGACATGAT	GCACAAATCC	ACTACGTGCA	CGCGATGAAG	180
GGCCGAGATTCA	CCATTCTTAG	ATGACAATGCC	AAGAATACAT	TGTATCTACA	ACTGAGCGGC	240
CTCAAAACCTG	AGGACACGGC	CATGTATTAC	TGTGCGCGG	TTGAATATCC	TATTGCAAGC	300
ATGTGTTTCA	GATACGGCGA	CCCGGGGACC	CAGGTCACCG	TCTCTCTACT	AGTTACCCGT	360
ACGACGAAAC	GGACTACGGT	TCTTAAATAGA	ATTG			394

( 2 ) INFORMATION FOR SEQ ID NO:104:

( i ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 433 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( 1 1 ) MOLECULE TYPE: cDNA

( x i ) SBOUENCE DESCRIPTION: SBO ID NO:104:

CTCGAGTCTG	GGGGAGGCTC	GGTCAGAGGT	GGAGGGTCTC	TGAAACTCTC	CTGTAAAAATC	60
TCTGGAGGTA	CCCCAGATCT	TGTTTCTAAA	CTTTTGGCCT	GGTTCGCGCA	GGCTCCAGAG	120
AAGGGACCGG	AGGGGATCTG	GATCTTTTTC	ACTAAGGATG	TGAAGCATTT	CTATCCCGAC	180
TCCTGTGAGG	CGCGATTAC	AGTTTCTTTA	GATATGACAA	GTAAACACTT	CTCTGTACAA	240
CTTGATCGAC	TGAACCCGGA	GGACACTGCC	GACTACTACT	CGCTGCAAA	TCAATTAGCT	300
GGTGCTGTGT	ATTTGAGACC	GAAATTACTG	CTCTCTGTGG	GTGCATATGC	CATCTGGGGC	360
CAGGGGACCC	AGGTACCGT	CTCCTCACT	GTTACCCGTA	CGACGTTTCG	GACTACGGTT	420
CTTAATAGTA	TTT					433

## ( 2 ) INFORMATION FOR SEQ ID NO:105:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 416 base pairs  
 ( B ) TYPE: nucleic acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:105:

CAGGTGAAAC	TGCTCGAGTC	TGGGGGAGGC	TCGGTGCAAG	CTGGGGGGTC	CTGACACTC	60
TCITGTGTAT	ACACCAACGA	TACTGGGACC	ATGGGATGGT	TTCGCCAAGC	TCCAGGGAAA	120
GAGTGCAGAA	GGGTCGCGCA	TATTACGCCT	GATGGTATGA	CCTTCATTGA	TGAACCCGTC	180
AAGGGGCGAT	TCACGATCTC	CCGAGACAAAC	GCCCCAGAAA	CGTTGTCTTT	GCGAATGAAT	240
AGTCTGAGGC	CTGAGGACAC	GCCCGTGAT	TACTGTGCGG	CAGATTGGAA	ATACTGGACT	300
TGTGGTGCCC	AGACTGGAGG	ATACTTCGGA	CAGTGGGGTC	AGGGGGCCCC	GGTCACCGTC	360
TCCTCACTAG	CTAGTTACCC	GTACGACGTT	CCGGACTACG	GTCTTAATA	GAATTTC	416

## ( 2 ) INFORMATION FOR SEQ ID NO:106:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 361 base pairs  
 ( B ) TYPE: nucleic acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:106:

CTCGAGTCTG	GGGGAGGCTC	GGTCCAACTT	GGAGGATCTC	TGACACTCTC	CTGTACAGTT	60
TCGGGGGCCA	CTACAGTGA	CTACAGTATT	GGATGGATCC	GCCAAGCTCC	AGGGAAAGAC	120
CGTGAAGTAG	TCGCAGCCGC	TAATACTGGT	GGGACTAGTA	AATTCTACGT	GCACTTTGTG	180
AAGGGCGGAT	TCACCATTTT	CCAAAGACAA	GCCAAAGATA	CGGTATATCT	GCAAAATGAGC	240
TTCTGAAAC	CTGAGGACAC	GCCCATCTAT	TACTGTGCGG	CAGCGGACCC	AAGTATATAT	300
TATAGTATCC	TCCATTGAGT	ATAAGTACTG	GGGCCAGGGG	ACCCAGGTCA	CCGTCTCCTC	360
A						361

## ( 2 ) INFORMATION FOR SEQ ID NO:107:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 354 base pairs  
 ( B ) TYPE: nucleic acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:107:

CTCGAGTCAO	GGGGAGGCTC	GGTGGAGGCT	GGAGGGTCTC	TGAGACTCTC	CTGTACAGCC	60
TCGGGATACG	TATCCTCTAT	GGCCTGGTTC	CGCCAGGGTC	CAGGGCAGGA	GCGCGAGGGG	120
GTCCGCTTTG	TTCAAACGCG	TGACAATAGT	GCATTATATG	GGCACTCCGT	GAAAGGCCGA	180
TTACCATCT	CCACAGACAA	GCACCAAGAAC	ACGGTGATCT	TGCAAAATGC	CAACCTGCAC	240
CCTGACGACA	CTGGCGTGTA	CTACTGTGCG	GCCCCAAAAG	AGGATCGTAC	TAGATGGGCC	300
GAGCCTCGAG	AATGGAACAA	CTGGGGCCAG	GGGACCCAGG	TCACCGTCTC	CTCA	354

-continued

## ( 2 ) INFORMATION FOR SEQ ID NO:108:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 381 base pairs  
 ( B ) TYPE: nucleic acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: cDNA

( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:108:

```

CTCGAGTCAG GTGTCCGGTC TGATGTGCAG CTGGTGGCGT CTGGGGGAGG CTCGGTGCAG      60
GCTGGAGGCT CTCTGAGACT CTCCTGTACA GCCTCTGGAG ACAATTTCAG TAGATTTCGC      120
ATGTCTTGGT TCCGCCAGGC TCCAGGGGAAO GAGTGCGAAT TGCTCTCAAG CATTCAAAGT      180
AATGGAAAGG CAACTGAGGC CGATTCCGTG CAAAGCCGAT TCACCATCTC CCGAGACAAT      240
TCCAAGGAAC CAGTGATCTC GCAAAAGAAC AGCCTGAAAC CCGAAGACAC GGCCTGTGAT      300
TACTGTGGGG CAGTCTCCCT AATGGACCGA ATTTCCTCAAC ATGGGTGCCG GGGCCAGGGA      360
ACCCAGGTCA CCGTCTCCTT A

```

## ( 2 ) INFORMATION FOR SEQ ID NO:109:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 18 amino acids  
 ( B ) TYPE: amino acid  
 ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: peptide

( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:109:

```

Gly  Gln  Pro  Arg  Glu  Pro  Gln  Val  Tyr  Thr  Leu  Pro  Pro  Ser  Arg  Asp
 1          5          10          15
Glu  Leu

```

## ( 2 ) INFORMATION FOR SEQ ID NO:110:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 18 amino acids  
 ( B ) TYPE: amino acid  
 ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: peptide

( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:110:

```

Gly  Gln  Pro  Arg  Glu  Pro  Gln  Val  Tyr  Thr  Leu  Pro  Pro  Ser  Arg  Glu
 1          5          10          15
Glu  Met

```

## ( 2 ) INFORMATION FOR SEQ ID NO:111:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 18 amino acids  
 ( B ) TYPE: amino acid  
 ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: peptide

( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:111:

```

Gly  Gln  Pro  Arg  Glu  Pro  Gln  Val  Tyr  Thr  Leu  Pro  Pro  Ser  Gln  Glu
 1          5          10          15
Glu  Met

```

## ( 2 ) INFORMATION FOR SEQ ID NO:112:



---

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 24 amino acids
      ( B ) TYPE: amino acid
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Leu  Leu  Gly  Gly  Pro  Ser  Val  Phe  Leu  Phe  Pro  Pro  Lys  Pro  Lys  Asp
1          5          10          15

Thr  Leu  Met  Ile  Ser  Arg  Thr  Pro
          20

( 2 ) INFORMATION FOR SEQ ID NO:113:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 23 amino acids
      ( B ) TYPE: amino acid
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:113:

Val  Ala  Gly  Pro  Ser  Val  Phe  Leu  Phe  Pro  Pro  Lys  Pro  Lys  Asp  Thr
1          5          10          15

Leu  Met  Ile  Ser  Arg  Thr  Pro
          20

( 2 ) INFORMATION FOR SEQ ID NO:114:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 24 amino acids
      ( B ) TYPE: amino acid
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:114:

Phe  Leu  Gly  Gly  Pro  Ser  Val  Phe  Leu  Phe  Pro  Pro  Lys  Pro  Lys  Asp
1          5          10          15

Thr  Leu  Met  Ile  Ser  Arg  Thr  Pro
          20

( 2 ) INFORMATION FOR SEQ ID NO:115:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 19 amino acids
      ( B ) TYPE: amino acid
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:115:

Lys  Gly  Glu  Pro  Arg  Glu  Pro  Glu  Val  Tyr  Thr  Leu  Pro  Pro  Ser  Arg
1          5          10          15

Asp  Glu  Leu

( 2 ) INFORMATION FOR SEQ ID NO:116:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 19 amino acids
      ( B ) TYPE: amino acid
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:116:

```

-continued

---

```

Lys Gly Glx Pro Arg Glu Pro Glx Val Tyr Thr Leu Pro Pro Ser Arg
1           5           10           15
Glu Glu Met

```

## (2) INFORMATION FOR SEQ ID NO:117:

```

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 19 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

```

(i) MOLECULE TYPE: peptide

(x) SEQUENCE DESCRIPTION: SEQ ID NO:117:

```

Lys Gly Glx Pro Arg Glu Pro Glx Val Tyr Thr Leu Pro Pro Ser Glx
1           5           10           15
Glu Glu Met

```

## (2) INFORMATION FOR SEQ ID NO:118:

```

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 30 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

```

(i) MOLECULE TYPE: peptide

(x) SEQUENCE DESCRIPTION: SEQ ID NO:118:

```

Glu Val Lys Leu Val Glu Ser Gly Gly Gly Leu Val Glu Pro Gly Gly
1           5           10           15
Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Phe Thr Phe Ser
                20           25           30

```

## (2) INFORMATION FOR SEQ ID NO:119:

```

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 29 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

```

(i) MOLECULE TYPE: peptide

(x) SEQUENCE DESCRIPTION: SEQ ID NO:119:

```

Glu Val Glx Leu Leu Ser Gly Gly Gly Leu Val Glx Pro Gly Gly Ser
1           5           10           15
Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser
                20           25

```

## (2) INFORMATION FOR SEQ ID NO:120:

```

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 11 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

```

(i) MOLECULE TYPE: peptide

(x) SEQUENCE DESCRIPTION: SEQ ID NO:120:

```

Trp Gly Glx Gly Thr Thr Val Thr Val Ser Ser
1           5           10

```

## (2) INFORMATION FOR SEQ ID NO:121:

```

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 11 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

```

-continued

---

```

( i i ) MOLECULE TYPE: peptide
( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:121:
  Trp  Gly  Gln  Gly  Thr  Met  Val  Thr  Val  Ser  Ser
   1          5          10

( 2 ) INFORMATION FOR SEQ ID NO:122:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amino acid
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide
( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:122:
  Trp  Gly  Gln  Gly  Thr  Thr  Leu  Thr  Val  Ser  Ser
   1          5          10

( 2 ) INFORMATION FOR SEQ ID NO:123:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amino acid
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide
( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:123:
  Trp  Gly  Gln  Gly  Thr  Ser  Val  Thr  Val  Ser  Ala
   1          5          10

( 2 ) INFORMATION FOR SEQ ID NO:124:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amino acid
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide
( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:124:
  Trp  Gly  Ala  Gly  Thr  Thr  Val  Thr  Val  Ser  Ser
   1          5          10

( 2 ) INFORMATION FOR SEQ ID NO:125:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 21 amino acids
      ( B ) TYPE: amino acid
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide
( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:125:
  Asp  Tyr  Tyr  Gly  Ser  Ser  Tyr  Phe  Asp  Val  Trp  Gly  Ala  Gly  Thr  Thr
   1          5          10          15
  Val  Thr  Val  Ser  Ser
           20

( 2 ) INFORMATION FOR SEQ ID NO:126:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 67 amino acids
      ( B ) TYPE: amino acid
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

```

-continued

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

```

Lys Val Asp Lys Arg Val Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr
1           5           10           15
His Thr Cys Pro Arg Cys Pro Glu Pro Lys Cys Ser Asp Thr Pro Pro
20
Pro Cys Pro Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro
35           40           45
Cys Pro Arg Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
50           55           60
Leu Phe Pro
65

```

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 35 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

```

Lys Val Asp Lys Lys Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr
1           5           10           15
Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
20           25           30
Leu Phe Pro
35

```

(2) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

```

Lys Val Lys Val Thr Val Glu Arg Lys Cys Cys Val Glu Cys Pro Pro
1           5           10           15
Cys Pro Ala Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro
20           25           30

```

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

```

Lys Val Asp Lys Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser
1           5           10           15
Cys Pro Ala Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro
20           25           30

```

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 11 amino acids

-continued

( B ) TYPE: amino acid  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Trp Gly Arg Gly Thr Leu Val Thr Val Ser Ser  
1 5 10

We claim:

1. A cDNA library comprising nucleotide sequences coding for a heavy-chain immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable ( $V_{HH}$ ) region and a constant region, said constant region being devoid of first constant domain  $C_{H1}$ , wherein said immunoglobulin is devoid of polypeptide light chains, obtained by performing the following steps:

- (a) obtaining B lymphocytes from a biological sample containing lymphoid cells, wherein said biological sample is obtained from a Camelid;
- (b) separating polyadenylated RNA from other nucleic acids and components of the B lymphocytes;
- (c) reacting the obtained RNA with a reverse transcriptase in order to obtain the corresponding cDNA;
- (d) contacting the obtained cDNA with 5' primers corresponding to mouse  $V_H$  domain of four-chain immunoglobulins, which primer contains a determined restriction site, and with 3' primers corresponding to the N-terminal part of a  $C_{H2}$  domain;
- (e) amplifying the DNA;
- (f) cloning the amplified DNA in a vector; and
- (g) recovering the clones hybridizing with a probe corresponding to the sequence coding for a constant domain of an isolated heavy-chain immunoglobulin.

2. A cDNA library comprising nucleotide sequences coding for a heavy-chain immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable ( $V_{HH}$ ) region and a constant region, said constant region being devoid of first constant domain  $C_{H1}$ , wherein said immunoglobulin is devoid of polypeptide light chains, obtained by performing the following steps:

- (a) obtaining B lymphocytes from a biological sample containing lymphoid cells, wherein said biological sample is obtained from a Camelid;
  - (b) separating polyadenylated RNA from other nucleic acids and components of the B lymphocytes;
  - (c) reacting the obtained RNA with a reverse transcriptase in order to obtain the corresponding cDNA;
  - (d) contacting the obtained cDNA with 5' primers located in the promoter, leader or framework sequences of the  $V_{HH}$  sequence of a heavy chain immunoglobulin, which primer contains a determined restriction site, and with the 3' primers located in the hinge,  $C_{H2}$ ,  $C_{H3}$ , 3' untranslated region or polyA tail;
  - (e) amplifying the DNA;
  - (f) cloning the amplified DNA in a vector; and
  - (g) recovering the obtained clones.
3. The cDNA library according to claim 2, wherein the B-lymphocytes of step (a) are obtained from an animal previously immunized against a determined antigen and the clones recovered in step (g) encode polypeptide chains having a preselected specificity for the antigen used for immunization.
4. The cDNA library according to claim 1 or claim 2, wherein said lymphoid cells are selected from the group consisting of peripheral lymphocytes, spleen cells, lymph nodes, and other lymphoid tissue.
5. The cDNA library according to claim 1 or claim 2, wherein in step (d) said restriction site is an XhoI site.
6. The cDNA library according to claim 1 or claim 2, wherein in step (f) said vector is a bluescript vector.

\* \* \* \* \*



US005840526A

**United States Patent** [19]**Casterman et al.**[11] **Patent Number:** **5,840,526**[45] **Date of Patent:** **Nov. 24, 1998**[54] **IMMUNOGLOBULINS DEVOID OF LIGHT CHAINS**[75] Inventors: **Cecile Casterman; Raymond Hamers**, both of Sint-Genesius-Rode, Belgium[73] Assignee: **Vrije Universiteit Brussel**, Brussels, Belgium[21] Appl. No.: **471,282**[22] Filed: **Jun. 6, 1995****Related U.S. Application Data**

[62] Division of Ser. No. 106,944, Aug. 17, 1993, abandoned.

[30] **Foreign Application Priority Data**Aug. 21, 1992 [EP] European Pat. Off. .... 92402326  
May 21, 1993 [EP] European Pat. Off. .... 93401310[51] **Int. Cl.<sup>6</sup>** ..... **C12P 21/06; C12N 15/63; C12N 5/04; C07K 16/12**[52] **U.S. Cl.** ..... **435/69.1; 435/320.1; 435/410; 830/387.1; 830/387.3; 830/388.22; 830/388.26; 830/388.4; 830/388.6; 830/391.7; 830/866**[58] **Field of Search** ..... **530/387.1, 387.3, 530/388.22, 388.26, 388.4, 388.6, 388.21, 391.7, 866, 867; 435/69.1, 320.1, 410****References Cited****FOREIGN PATENT DOCUMENTS**

WO 92/01787 2/1992 WIPO.

**OTHER PUBLICATIONS**Blier et al., "A Limited Number of BfCell Lineages Generates The Heterogeneity Of A Secondary Immune Response", *J. Immunology*, vol. 139, No. 12, pp. 3996-4006 (Dec. 15, 1987).

Websters II New Riverside University Dictionary, Riverside Publishing Co./Houghton Mifflin Co., Boston, Ma., 1995.

Hamers—Casterman et al., "Nature", 363:466-48 (1993).

Roitt et al., "Immunology", Gower Medical Publishing, London, pp. 1.5 and 5.7 (1985).

Ward et al., "Nature", 341:544-546 (1989).

*Primary Examiner*—Frank C. Eiscschenk*Assistant Examiner*—Evelyn Rabin*Attorney, Agent, or Firm*—Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.

[57]

**ABSTRACT**

There is provided an isolated immunoglobulin comprising two heavy polypeptide chains sufficient for the formation of a complete antigen binding site or several antigen binding sites, wherein the immunoglobulin is further devoid of light polypeptide chains.

**10 Claims, 12 Drawing Sheets**

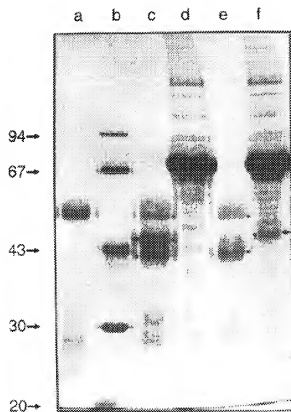


FIG. 1A

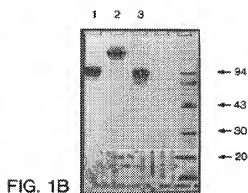


FIG. 1B

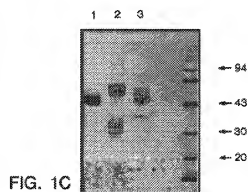


FIG. 1C

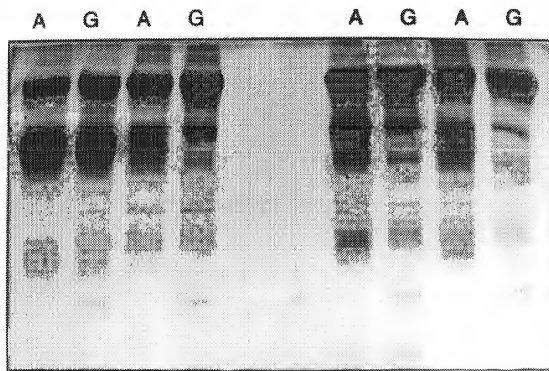


FIG. 2A

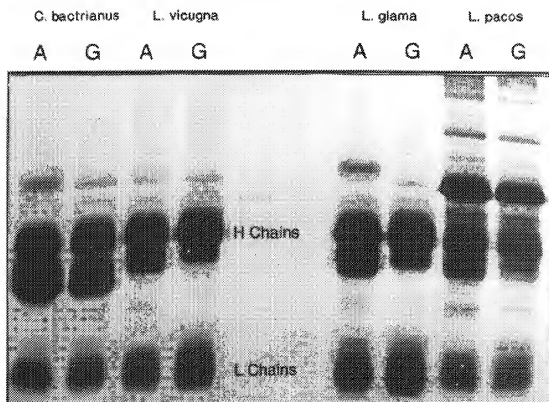


FIG. 2B



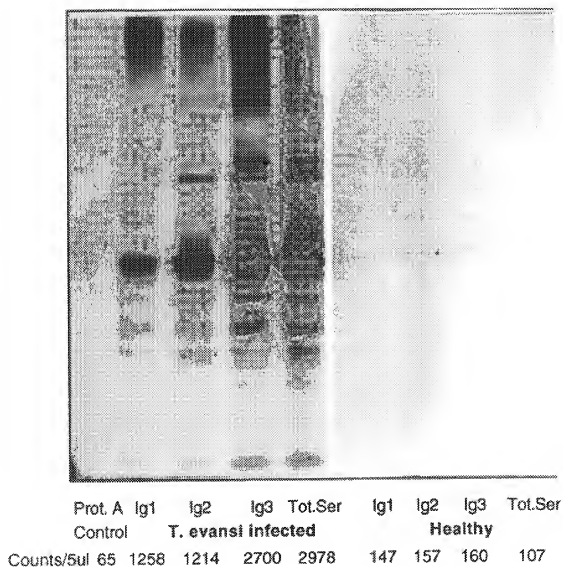


FIG. 3A

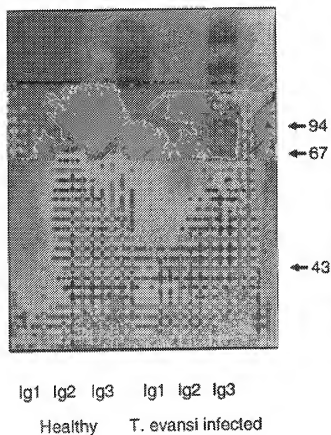


FIG. 3B

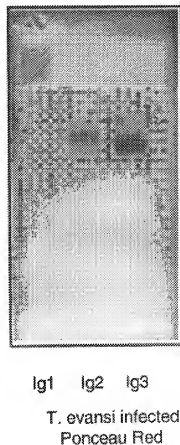


FIG. 3C

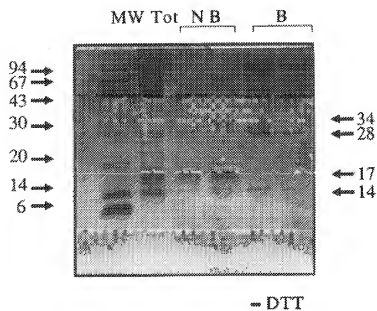


FIG. 4A

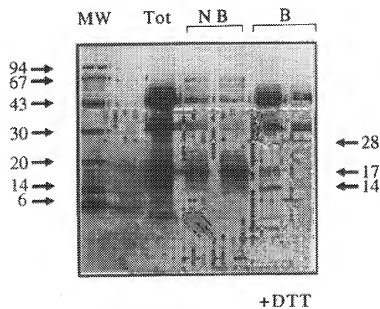


FIG. 4B

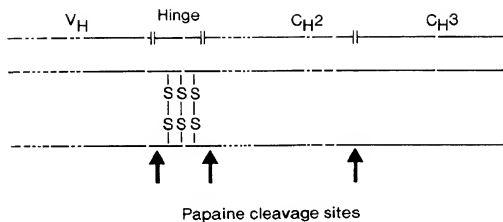


FIG. 5

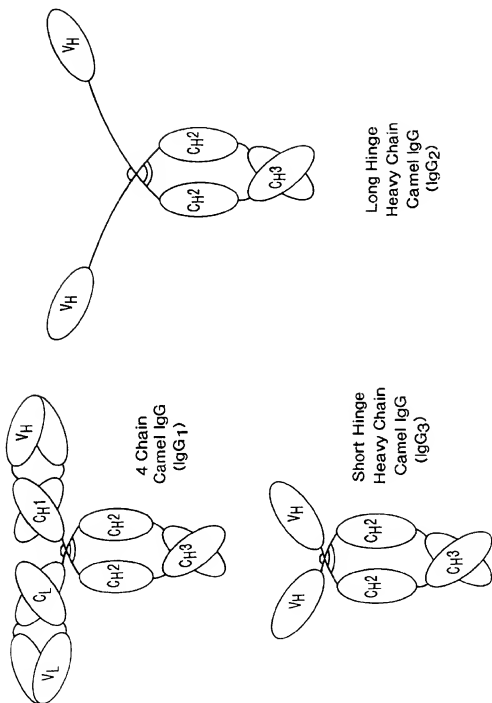


FIG. 6

```

DR01005 C-----TCGAG--TCTG6GGGAGG
DR27005 C-----TCGAG--TCTG6GGGAGG
DR03005 C-----AGGTGA-----AACTGCTCGAG--TCTG6AGGAGG
DR11005 C-----TCGAG--TCTG6GGGAGG
DR24005 C-----AGGTGA-----AACTGCTCGAG--TCTG6GGGAGG
DR16005 C-----TCGAG--TCTG6AGGAGG
DR19005 C-----TCGAG--TCTG6AGGAGG
DR07005 C-----TCGAG--TCTG6GGGAGG
DR16005 C-----TCGAG--TCTG6GGGAGG
DR20005 C-----TCGAG--TCA6GGGAGG
DR25005 C-----TCGAG--TCTG6GGGAGG
DR20005 C-----TCGAG--TCTG6AGGAGG
DR21005 C-----TCGAG--TCTG6GGGAGG
DR09005 C-----AGGTGA-----AACTGCTCGAG--TCTG6GGGAGG
DR17005 C-----TCGAG--TCTG6GGGAGG
DR13005 C-----TCGAG--TCA6GGGAGG
DR02005 CTCGAGTCA6GTGTCCGGTCTGATGTGCACTG6TGGCGTCTG6GGGAGG

```

```

DR01005 ATCGGTGCAGGCTGGAGGGTCTCTGAGACTCTC--GTGCG-CAGCCTCTG
DR27005 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTGCATCTTCTCTA
DR03005 CTCGGTGCAGACTGGAGGATCTCTGAGACTCTCTGTGCAGT--C-TCTG
DR11005 GTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTAATGT--C-TCTG
DR24005 GTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTAATGT--C-TCTG
DR16005 CTCGGCGCAGGCTGGAGGATCTCTGAGACTCTCTGTGCAGC--CCACGG
DR19005 CTCGGTTCAGGCTGGAGGGTCCCTTAGACTCTCTGTGCAGC--C-TCTG
DR07005 CTCGGTGCAGGGTGGAGGGTCTCTGAGACTCTCTGTGCAA--TCTCTG
DR16005 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTACAG--GCTCTG
DR20005 CTCGGTACAGGTTGGAGGGTCTCTGAGACTCTCTGTGTAG--CCTCTA
DR25005 CTCGGTACAACTGGAGGGTCTCTGAGACTCTCTTGC--AAATCTCTG
DR20005 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTG--TAGCCTCTG
DR21005 CTCGGTGCAGGTTGGAGGGTCTCTGAACTCTCTGTAAAA--CTCTG
DR09005 CTCGGTGCAGGCTGGGGGGTCTCTGACACTCTCTTGTG--TATACAG--
DR17005 CTCGGTCCAACCTGGAGGATCTCTGACACTCTCTGTACAGT--TCTG
DR13005 CTCGGTGGAGGCTGGAGGGTCTCTGAGACTCTCTGTACAG--CCTCTG
DR02005 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCTGTACAG--CCTCTG

```

```

DR01005 GA--TACAGTAATT---GTCCCTCACTTG-GAGCTGGTATCGCCAGTTT
DR27005 AA--TATATGCCTT---GCACCTACGACAT-GACCTGGTACCGCCAGGCT
DR03005 GA--TTCTCCTTTA---GTACAGTTGTAT-GGCTGGTCCGCCAGGCT
DR11005 GC--TCTCCAGTA---GTACTTATTGCCT-GGGCTGGTCCGCCAGGCT
DR24005 GC--TCTCCAGTA---GTACTTATTGCCT-GGGCTGGTCCGCCAGGCT
DR16005 GA--TTCCGC-TCA---ATGCTTACTACAT-CGCTGGTTCGGTCAAGGCT
DR19005 AC--TACACCATCA---CTGATTATTGCAT-GGCCTGGTCCGCCAGGCT
DR07005 GA--TACACGTACG---GTAGCTTCTGTAT-GGGCTGGTCCGCCAGGCT
DR16005 GA--TTCCCTATA---GTACCTTCTGTCT-GGGGTGGTCCGCCAGGCT
DR20005 CT--CACACCGACA---GTAGCACCTGTAT-AGGCTGGTCCGCCAGGCT
DR25005 GA--TTGACTTTTG---ATGATTCTGACGT-GGGGTGGTACCGCCAGGCT
DR20005 GA--TTCAATTTTCG---AACTTCTCGTAT-GGCCTGGTACCGCCAGGCT
DR21005 GAGGTACCCAGATCGTGTCTCTAAATCTTTGGCTGTGGTCCGCCAGGCT
DR09005 -----CAACGATACTGGGACCA-----TGGGATGGTTCCGCCAGGCT
DR17005 --GGGCACCTACA---GTGACTACAGTATGGA-TGGATCCGCCAGGCT
DR13005 G-----ATACGTAT-CCT---CTATGGCCTGGTCCGCCAGGCT
DR02005 GAGA---CAGTTTCAGTAGATT--TGCCATGTCTTGGTTCGCCAGGCT

```

FIG. 7A

DR01006 CCAGGAACGGAGCGCGAGTTCTCGTCTCCAGTATGGATCCGGATGGAAATAC  
DR27006 CCAGGCAAGGAGCGCGAATTTGTCTCAAGTATAAATATTGATGGTAAGAC  
DR03006 TCAGGAAGCAGCGTGAGGGGGTCCGAGCCATTAATAGTGGCGGTGGTAAG  
DR11006 CCAGGGAAGGAGCGTGAGGGGGTCAACGCGATTAA-----CACTGATGG  
DR24006 CCAGGGAAGGAGCGTGAGGGGGTCAACGCGATTAA-----CACTGATGG  
DR16006 CCTGGGAAGGGGCGTGAGGGGGTCCGAACAATTAATGGTGGTGG-----  
DR19006 CCAGGGAAGGAGCGTGGAATTGGTCCGAGCGATTCAAGTTGTCGGTAGTGA  
DR07006 CCAGGCAAGGAGCGTGAGGGGGATCGCAACTATTCTTAATGGTGGTACTAA  
DR16006 CCAGGGAAGGAGCGTGAGGGGGTCCGCGGTATTAATAGTGCAGGAGGTAA  
DR20006 CCAGGGAAGGAGCGCGAGGGGGTCCGCAAGTATATATTTTGGTGGTGGTGG  
DR25006 CCAGGGCATGAGTGCAAAATTGGTCTCAGGTATTCGAGTGATGGTACT-C  
DR20006 CCAGGAAATGTGTGTGAGTTGGTCTCAGGTATTTACAGTGATGG-----  
DR21006 CCAGAGAAGGAGCGCGAGGGGATCGCAGTTCTTTCCAGTAAGGATGGTAA  
DR09006 CCAGGGAAGGAGTCCGAAGGGGTCGCGCATATTACGCTGATGGTATGA-  
DR17006 CCAGGGAAGGAGCGTGGAAGTAGTCGAGCGCTAATACTGGTGG-----  
DR13006 CCAGGGCAGGAGCGCGAGGGGGTCCGTTTGTTCAAACGG-----  
DR02006 CCAGGGAAGGAGTGCGCAATTGGTCTCAGCATTCAAAGTAATGGAAGGAC

DR01006 CAAGTACA-----CATACTCCGTGAAGGGCCGCTTCACC  
DR27006 AACATACG-----CAGACTCCGTGAAGGGCCGATTCAAC  
DR03006 GACATACTA-CAACACATATGTCGCCGAGTCCGTGAAGGGCCGATTCCGC  
DR11006 CAGTATCAT-ATACGCA-----GCCGACTCCGTGAAGGGCCGATTCAAC  
DR24006 CAGTGTAT-ATACGCA-----GCCGACTCCGTGAAGGGCCGATTCAAC  
DR16006 -----CGA-CGTACATACTACGCCGACTCCGTGAAGGGCCGATTCAAC  
DR19006 TACT-----CGC-C-TCACAGACTACGCCGACTCCGTGAAGGGCCGATTCAAC  
DR07006 -----CACATACTATGCCGACTCCGTGAAGGGCCGATTCAAC  
DR16006 -----TACTTACTATGCCGACTCCGTGAAGGGCCGATTCAAC  
DR20006 -----TACGAATTATGCCGACTCCGTGAAGGGCCGATTCAAC  
DR25006 CATATACAAAGAGTGGAGACTATGCTGAGTCTGTGAGGGGCCGGGTACCC  
DR20006 CA-AAACATACTACGTCGACC--GCA-----TGAAGGGCCGATTCAAC  
DR21006 GA-----CATTCATGCCGACTCCGTGAAGGGCCGATTCAAC  
DR09006 -----CCTTCATTGATGAACCCGTGAAGGGCCGATTCAACG  
DR17006 -----CGACTAGTAAATTTACGTCGACTTTGTGAAGGGCCGATTCAAC  
DR13006 -----CTGCAAT-AGTGCAATTATATGGCGACTCCGTGAAGGGCCGATTCAAC  
DR02006 AACTGA-----GGCCGATTCCGTGAAGGGCCGATTCAAC

DR01006 ATGTCCCGAGGAGCACCAGTACACAGTATTTCTGCAAAATGGACAATCT  
DR27006 ATCTCCCAAGACAGCGCCAAGAACACGGTGTATCTGCAGATGAACAGCCT  
DR03006 ATCTCCCAAGACAACGCCAAGAACACCGGTATATCTTGATATGAACAACCT  
DR11006 ATCTCCCAAGACACCGCCAAGGAACCGGTACATCTCCAGATGAACAACCT  
DR24006 ATCTCCCAAGACACCGCCAAGGAACCGGTATATCTCCAGATGAACAACCT  
DR16006 ATCTCCCGAGACAGCCCCAAGAATAACGGTGTATCTGCAGATGAACAGCCT  
DR19006 ATCTCCCAAGGCAACACCAAGAACACAGTGAATCTGCAAAATGAACAGCCT  
DR07006 ATCTCCCAAGACAGCACGTTGAAGACGATGTATCTGCTAATGAACAACCT  
DR16006 ATCTCCCAAGGGAATGCCAAGAATAACGGTGTCTGCAAAATGGATAACTT  
DR20006 ATCTCCCAACTCAACGCCCGACAACACAGTGTATCTGCAAAATGAACAGCCT  
DR25006 ATCTCCAGAGACAACGCCAAGAACATGATATACCTTCAAATGAACAGCCT  
DR20006 ATTTCTAGAGAGAAATGCCAAGAATACATTGTATCTACAACTGAGCGGCT  
DR12006 ATCTTCTTAGATAAATGACAAGACCACTTTCTCTTCAACCTTGTACGACT  
DR09006 ATCTCCCGAGACAACGCCCGAAGAACGTTGTCTTTGCGAATGAATAGTCT  
DR17006 ATTTCCCAAGACAACGCCAAGAATAACGGTGTATCTGCAAAATGAGCTTCT  
DR13006 ATCTCCACGACAACGCCAAGAACACGCTGTATCTGCAAAATGCGCAACCT  
DR02006 ATCTCCCGAGACAATTCAGGAACACAGTGTATCTGCAAAATGAACAGCCT

FIG. 7B

DR01006 GAAACCTGAGGACACGGCGATGTATTACTGTAAAC-A---GCCCTAC--  
DR27006 GAAACCTGAGGACACGGCGATGTATTACTGTAAAC-A---GA---TTC--  
DR03006 AACCCCTGAAGACACGGCTACGTATTACTGTGCGGCGG---TCCAGGCC  
DR11006 CCAACCTGAGGATACGGCCACCTATTACTGCGCGGCAA---GACTGACGG  
DR24006 GCAACCTGAGGATACGGCCACCTATTACTGCGCGGCAA---GACTGACGG  
DR16006 GAAACCTGAGGACACGGCCATCTACTTCTGTGACGACG---G---CTC  
DR19006 GACACCTGAGGACACGGCCATCTACAGTTGTGCGGCAA---C---CAG  
DR07006 GAAACCTGAAGACACGGGCACCTATTACTGTGCTG-C---GAACTAAGT  
DR16006 GAAACCTGAGGACACGGCCATCTATTACTGCGCGG-CG---GATAGTCCA  
DR20006 GAAACCTGAGGACACGGCCATGTACTACTGTGCAATCA---CTGAAATTG  
DR25006 GAAACCTGAGGACACGGCCATGTATTACTGCGCGGTAGATGGTTGGACCC  
DR20006 CAAACCTGAGGACACGGCCATGTATTACTGTGCG---CC  
DR21006 GAACCTGAGGACATGCGGACTACTACTGCGCTGCAAAATCAATTAGC--  
DR09006 GAGGCTGAGGACACGGCCGTGTATTACTGTGCGGCAAGTTG---  
DR17006 GAAACCTGAGGACACGGCCATCTATTACTGTGCGGCGA---CGGACCC  
DR13006 CCAACCTGACGACACTGGCGTGTACTACTGTGCGGCC---CAA  
DR02006 GAAACCCGAGGACACGGCGTGTATTACTGTGGGCGAGT-----

DR01006 -----A-AC---CTGGGGGTATTGTGGGTA-  
DR27006 -----GTAC---CCGTGCCATCTCCTTGATG-  
DR03006 ACCTTGGGACCT-----GGCG-CAATT---CTGGATTG  
DR11006 AGATGGGGGCTTGTGATGCGAGATGGGCGACCTTAGC---GACAAGGAC-G  
DR24006 AGATGGGGGCTTGTGATGCGAGATGGGCGACCTTAGC---GACAAGGAC-G  
DR16006 GCGTTTTT-CTAGTCTGTGGGAGCACTTC-TAGAC---TCGAAAGTAG  
DR19006 TAGTTTTTACTGGTACT-----GCAC-----C---ACG-----G  
DR07006 GGTGGTAGTTGTGAATTGC---CTTTG-----TATTGTACA-----  
DR16006 TGTACATGCCGACTATGC---CCGCTCCCCGATACGAGACAGTTTTGG  
DR20006 AGTGGTATGGGTGCAATTT---AAGGACTACTTTTACT---C-----G  
DR25006 GGAAGGAAG---GGGAATCGGGTTAC---CCTGGTCGGTCAATGTGAA  
DR20006 GGTGAA-----TATC---CTATTGCAGAC---ATGTGTT  
DR21006 ---TGGTGGCTGGTATT-----TGGACCCGAATTACTGG---CTCTGTG  
DR09006 ---GAAATACTGGA---CTTGTGGTGC---CCAGA-CTGG-----AG  
DR17006 AAGTATATATTATAGTATC-----CTCCNNAT-----  
DR13006 AAGAAGGATCGTA-----CTAGATGGGC-----CGAGCCT-----  
DR02006 -----CTCCCTAA---TGGACCGAATTTCT

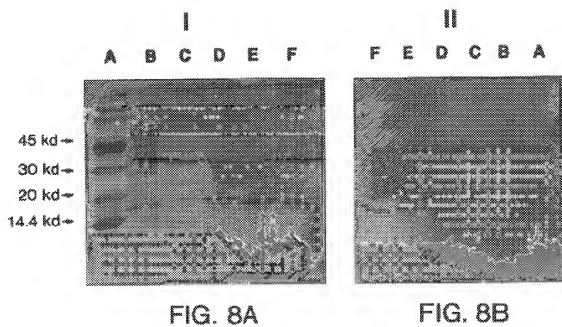
DR01006 --TGGGTANTGCCTCTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCACT  
DR27006 --T-----CTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCACT  
DR03006 AAAAAAGTATAAGTACTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCACT  
DR11006 TTTCGTATAACTACTGGGGCCGGGGGACCCAGGTCACCGTCTCCTCACT  
DR24006 TTTGCGTATAACTACTGGGGCCGGGGGACCCAGGTCACCGTCTCCTCACT  
DR16006 CGA-CT-ATAACTATTGGGGCCAGGGGATCCAGGTCACCGTCTCCTCACT  
DR19006 CGC-CTTATAACGTCCTGGGGTCAAGGGGACCCAGGTCACCGTCTCCTCACT  
DR07006 CTGGG-----GCCAGGGGACCCAGGTCACCGTCTCCTCACT  
DR16006 CTGGGATGATTIT-----GGCCAGGGGACCCAGGTCACCGTCTCCTCACT  
DR20006 CTGGG-----GCCAGGGGACCCAGGTCACCGTCTCCTCACT  
DR25006 GATGGTTATAACTATTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCACT  
DR20006 CGAGAT-----ACG---GCGACCCGGGGACCCAGGTCACCGTCTCCTCACT  
DR21006 GGTGCATATGCCATCTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCACT  
DR09006 GATACTTCGGACAG-TGGGGTCAAGGGGACCCAGGTCACCGTCTCCTCACT  
DR17006 --TGAGTATAAGTACTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCACT  
DR13006 CGAGAAATGGAACACTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCACT  
DR02006 CCAACATGGG--TGCCGGGGCCAGGGGACCCAGGTCACCGTCTCCTCT--

FIG. 7C



DR01006 AG----TTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR27006 AG----TTACCCGTACGAGCTTCCGGACTACGGTTCCTTAATAGAATTC  
DR03006 AGCTAGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR11006 AG----TTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR24006 AGCTAGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR16006 ----AGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR19006 ----AGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR07006 ----AGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR16006 ----AGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR20006 ----AGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR25006 ---TAGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR20006 ---TAGTTACCCGTACGACGAACCGGACTACGGTTCCTTAATAGAATTC  
DR21006 ---TAGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR09006 AGCTAGTTACCCGTACGACGTTCCGGACTACGGTTCCTTAATAGAATTC  
DR17006 -----  
DR13006 -----  
DR02006 -----TA

FIG. 7D



# IMMUNOGLOBULINS DEVOID OF LIGHT CHAINS

This is a division of application Ser. No. 08/106,944, filed Aug. 17, 1993, now abandoned.

The invention relates to new isolated immunoglobulins which are devoid of light polypeptide chains. These immunoglobulins do not consist in the degradation products of immunoglobulins composed of both heavy polypeptide and light polypeptide chains but to the contrary, the invention defines a new member of the family of the immunoglobulins, especially a new type of molecules capable of being involved in the immune recognition. Such immunoglobulins can be used for several purposes, especially for diagnosis or therapeutical purposes including protection against pathological agents or regulation of the expression or activity of proteins.

Up to now the structure proposed for immunoglobulins consists of a four-chain model referring to the presence of two identical light polypeptide chains (light chains) and two identical heavy polypeptide chains (heavy chains) linked together by disulfide bonds to form a Y- or T-shaped macromolecules. These chains are composed of a constant region and a variable region, the constant region being subdivided in several domains. The two heavy polypeptide chains are usually linked by disulfide bounds in a so-called "hinge region" situated between the first and second domains of the constant region.

Among the proteins forming the class of the immunoglobulins, most of them are antibodies and accordingly present an antigen binding site or several antigen binding sites.

According to the four-chain model, the antigen binding site of an antibody is located in the variable domains of each of the heavy and light chains, and requires the association of the heavy and the light chains variable domains.

For the definition of these four-chain model immunoglobulins, reference is made to Roitt, I. et al (Immunology-second-Edition Gower Medical Publishing U.S.A., 1989). Reference is especially made to the part concerning the definition of the four-chain immunoglobulins, their polypeptidic and genetic structures, the definition of their variable and constant regions and the obtention of the fragments produced by enzymatic degradation according to well known techniques.

The inventors have surprisingly established that different molecules can be isolated from animals which naturally produce them, which molecules have functional properties of immunoglobulins these functions being in some cases related to structural elements which are distinct from those involved in the function of four-chain immunoglobulins due for instance to the absence of light chains.

The invention relates to two-chain model immunoglobulins which neither correspond to fragments obtained for instance by the degradation in particular the enzymatic degradation of a natural four-chain model immunoglobulin, nor correspond to the expression in host cells, of DNA coding for the constant or the variable region of a natural four-chain model immunoglobulin or a part of these regions, nor correspond to antibodies produced in lymphocytes for example in mice, rats or human.

E. S. Ward et al (1) have described some experiments performed on variable domains of heavy polypeptide chains ( $V_H$ ) or/and light polypeptide chains ( $V_L$ ) to test the ability of these variable domains, to bind specific antigens. For this purpose, a library of  $V_H$  genes was prepared from the spleen genomic DNA of ice previously immunized with these specific antigens.

Ward et al have described in their publication that  $V_H$  domains are relatively sticky, presumably due to the exposed hydrophobic surface normally capped by the  $V_K$  or  $V_L$  domains. They consequently envisage that it should be possible to design  $V_H$  domains having improved properties and further that  $V_H$  domains with binding activities could serve as the building blocks for making variable fragments (Fv fragments) or complete antibodies.

The invention does not start from the idea that the different fragments (light and heavy chains) and the different domains of these fragments of four-chain model immunoglobulin can be modified to define new or improved antigen binding sites or a four-chain model immunoglobulin.

The inventors have determined that immunoglobulins can have a different structure than the known four-chain model and that such different immunoglobulins offer new means for the preparation of diagnosis reagents, therapeutical agents or any other reagent for use in research or industrial purposes.

Thus the invention provides new immunoglobulins which are capable of showing functional properties of four-chain model immunoglobulins although their structure appears to be more appropriate in many circumstances for their use, their preparation and in some cases for their modification. Moreover these molecules can be considered as lead structures for the modification of other immunoglobulins. The advantages which are provided by these immunoglobulins comprise the possibility to prepare them with an increased facility.

The invention accordingly relates to immunoglobulins characterized in that they comprise two heavy polypeptide chains sufficient for the formation of a complete antigen binding site or several antigen binding sites, these immunoglobulins being further devoid of light polypeptide chains. In a particular embodiment of the invention, these immunoglobulins are further characterized by the fact that they are the product of the expression in a prokaryotic or in a eukaryotic host cell, of a DNA or of a cDNA having the sequence of an immunoglobulin devoid of light chains as obtainable from lymphocytes or other cells of Camelids.

The immunoglobulins of the invention can be obtained for example from the sequences which are described in FIG. 7.

The immunoglobulins of the invention, which are devoid of light chains are such that the variable domains of their heavy chains have properties differing from those of the four-chain immunoglobulin  $V_H$ . The variable domain of a heavy-chain immunoglobulin of the invention has no normal interaction sites with the  $V_L$  or with the  $C_H1$  domain which do not exist in the heavy chain immunoglobulins. It is hence a novel fragment in many of its properties such as solubility and position of the binding site. For clarity reasons we will call it  $V_{Hf}$  in this text to distinguish it from the classical  $V_H$  of four-chain immunoglobulins.

By "a complete antigen binding site" it is meant according to the invention, a site which will alone allow the recognition and complete binding of an antigen. This could be verified by any known method regarding the testing of the binding affinity.

These immunoglobulins which can be prepared by the technique of recombinant DNA, or isolated from animals, will be sometimes called "heavy-chain immunoglobulins" in the following pages. In a preferred embodiment of the invention, these immunoglobulins are in a pure form.

In a first embodiment, the immunoglobulins of the invention are obtainable in prokaryotic cells, especially in *E. coli* cells by a process comprising the steps of:

a) cloning in a Bluescript vector of a DNA or cDNA sequence coding for the  $V_H$  domain of an immunoglobulin devoid of light chain obtainable for instance from lymphocytes of Camelids,

b) recovering the cloned fragment after amplification using a 5' primer containing an Xho site and a 3' primer containing the Spe site having the following sequence TC TTA ACT AGT GAG GAG ACG GTG ACC TG, SEQ ID NO: 51

c) cloning the recovered fragment in phase in the immuno PBS vector after digestion of the vector with Xho and Spe restriction enzymes,

d) transforming host cells, especially *E. coli* by transfection with the recombinant immuno PBS vector of step c,

e) recovering the expression product of the  $V_H$  coding sequence, for instance by using antibodies raised against the dromedary  $V_H$  domain.

In another embodiment the immunoglobulins are hetero-specific immunoglobulins obtainable by a process comprising the steps of:

obtaining a first DNA or cDNA sequence coding for a  $V_H$  domain or part thereof having a determined specificity against a given antigen and comprised between Xho and Spe sites,

obtaining a second DNA or cDNA sequence coding for a  $V_H$  domain or part thereof, having a determined specificity different from the specificity of the first DNA or cDNA sequence and comprised between the Spe and EcoRI sites,

digesting an immuno PBS vector with EcoRI and XhoI restriction enzymes,

ligating the obtained DNA or cDNA sequences coding for  $V_H$  domains, so that the DNA or cDNA sequences are serially cloned in the vector,

transforming a host cell, especially *E. coli* cell by transfection, and recovering the obtained immunoglobulins.

In another embodiment, the immunoglobulins are obtainable by a process comprising the steps of:

obtaining a DNA or cDNA sequence coding for a  $V_H$  domain or part thereof, having a determined specific antigen binding site,

amplifying the obtained DNA or cDNA, using a 5' primer containing an initiation codon and a HindIII site, and a 3' primer containing a termination codon having a XhoI site,

recombining the amplified DNA or cDNA into the HindIII (position 2650) and XhoI (position 4067) sites of a plasmid pMM984,

transfecting permissive cells especially NB-E cells with the recombinant plasmid,

recovering the obtained products.

Successful expression can be verified with antibodies directed against a region of a  $V_H$  domain, especially by an ELISA assay.

According to another particular embodiment of this process, the immunoglobulins are cloned in a parvovirus. In another example these immunoglobulins are obtainable by a process comprising the further cloning of a second DNA or cDNA sequence having another determined antigen binding site, in the pMM984 plasmid.

Such an Immunoglobulin can be further characterized in that it is obtainable by a process wherein the vector is Yep 52 and the transformed recombinant cell is a yeast especially *S. cerevisiae*.

A particular Immunoglobulin is characterized in that it has a catalytic activity, especially in that it is directed against an antigen mimicking an activated state of a given substrate. These catalytic antibodies can be modified at the level of their binding site, by random or directed mutagenesis in order to increase or modify their catalytic function. Reference may be made to the publication of Lerner et al (TIBS November 1987, 427-430) for the general technique for the preparation of such catalytic immunoglobulins.

According to a preferred embodiment, the immunoglobulins of the invention are characterized in that their variable regions contain in position 45, an amino-acid which is different from leucine, proline or glutamine residue.

Moreover the heavy-chain immunoglobulins are not products characteristic of lymphocytes of animals nor from lymphocytes of a human patient suffering from lymphopathies. Such immunoglobulins produced in lymphopathies are monoclonal in origin and result from pathogenic mutations at the genomic level. They have apparently no antigen binding site.

The two heavy polypeptide chains of these immunoglobulins can be linked by a hinge region according to the definition of Roitt et al.

In a particular embodiment of the invention, immunoglobulins corresponding to the above-defined molecules are capable of acting as antibodies.

The antigen binding site(s) of the immunoglobulins of the invention are located in the variable region of the heavy chain.

In a particular group of these immunoglobulins each heavy polypeptide chain contains one antigen binding site on its variable region, and these sites correspond to the same amino-acid sequence.

In a further embodiment of the invention the immunoglobulins are characterized in that their heavy polypeptide chains contain a variable region ( $V_H$ ) and a constant region ( $C_H$ ) according to the definition of Roitt et al, but are devoid of the first domain of their constant region. This first domain of the constant region is called  $C_{H1}$ .

These immunoglobulins having no  $C_{H1}$  domain are such that the variable region of their chains is directly linked to the hinge region at the C-terminal part of the variable region.

The immunoglobulins of the type described here-above can comprise type G immunoglobulins and especially immunoglobulins which are defined as immunoglobulins of class 2 (IgG2) or immunoglobulins of class 3 (IgG3).

The absence of the light chain and of the first constant domain lead to a modification of the nomenclature of the immunoglobulin fragments obtained by enzymatic digestion, according to Roitt et al.

The terms Fc and pFc on the one hand, Fe' and pFe' on the other hand corresponding respectively to the papain and pepsin digestion fragments are maintained.

The terms Fab F(ab), F(ab)', Fabe, Fd and Fv are no longer applicable in their original sense as these fragments have either a light chain, the variable part of the light chain or the  $C_{H1}$  domain.

The fragments obtained by papain digestion and composed of the  $V_H$  domain and the hinge region will be called FV<sub>H</sub>h or F(V<sub>H</sub>h)<sub>2</sub> depending upon whether or not they remain linked by the disulphide bonds.

In another embodiment of the invention, immunoglobulins replying to the heretofore given definitions can be originating from animals especially from animals of the camelid family. The inventors have found out that the heavy-chain immunoglobulins which are present in camelids are not associated with a pathological situation which

would induce the production of abnormal antibodies with respect to the four-chain immunoglobulins. On the basis of a comparative study of old world camelids (*Camelus bactrianus* and *Camelus dromaderius*) and new world camelids (for example *Lama paccos*, *Lama glama*, and *Lama vicugna*) the inventors have shown that the immunoglobulins of the invention, which are devoid of light polypeptide chains are found in all species. Nevertheless differences may be apparent in molecular weight of these immunoglobulins depending on the animals. Especially the molecular weight of a heavy chain contained in these immunoglobulins can be from approximately 43 kd to approximately 47 kd, in particular 45 kd.

Advantageously the heavy-chain immunoglobulins of the invention are secreted in blood of camelids.

Immunoglobulins according to this particular embodiment of the invention are obtainable by purification from serum of camelids and a process for the purification is described in details in the examples. In the case where the immunoglobulins are obtained from Camelids, the invention relates to immunoglobulins which are not in their natural biological environment.

According to the invention immunoglobulin IgG2 as obtainable by purification from the serum of camelids can be characterized in that:

it is not adsorbed by chromatography on Protein G Sepharose column,

it is adsorbed by chromatography on Protein A Sepharose column,

it has a molecular weight of around 100 kd after elution with a pH 4.5 buffer (0.15M NaCl, 0.58% acetic acid adjusted to pH 4.5 by NaOH),

it consists of heavy  $\beta$ 2 polypeptide chains of a molecular weight of around 46 kd preferably 45 after reduction.

According to a further embodiment of the invention another group of immunoglobulins corresponding to IgG3, as obtainable by purification from the serum of Camelids is characterized in that the immunoglobulin:

is adsorbed by chromatography on a Protein A Sepharose column,

has a molecular weight of around 100 kd after elution with a pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid),

is adsorbed by chromatography on a Protein G Sepharose column and eluted with pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid),

consists of heavy  $\gamma$ 3 polypeptide chains of a molecular weight of around 45 kd in particular between 43 and 47 kd after reduction.

The immunoglobulins of the invention which are devoid of light chains, nevertheless comprise on their heavy chains a constant region and a variable region. The constant region comprises different domains.

The variable region of immunoglobulins of the invention comprises frameworks (FW) and complementarity determining regions (CDR), especially 4 frameworks and 3 complementarity regions. It is distinguished from the four-chain immunoglobulins especially by the fact that this variable region can itself contain an antigen binding site or several, without contribution of the variable region of a light chain which is absent.

The amino-acid sequences of frameworks 1 and 4 comprise among others respectively amino-acid sequences which can be selected from the following for the framework 1 domain

---

```
GGSVQTGGSLRLSCSEISGLTFDSEQ ID NO:1
GGSVQTGGSLRLSCAVSGFSFSEQ ID NO:2
GGSEGGGSLRLSCAISGYTYGSEQ ID NO:3
GGSVQPGGSLTSLCTVSGATYSSEQ ID NO:4
GGSVQAGGSLRLSCTGSGFFYSSEQ ID NO:5
GGSVQAGGSLRLSCTVAGFGTTSSEQ ID NO:6
GGSVQAGGSLRLSCTVSPFSSEQ ID NO:7
```

---

10 for the framework 4 domain

---

```
WGGGTQVTVSSSEQ ID NO:8
WGGTTLTVTVSSSEQ ID NO:9
WGGGAQVTVSSSEQ ID NO:10
WGGGTQVTASSSEQ ID NO:11
RGGGTQVTVTVSSSEQ ID NO:12
```

---

for the CDR3 domain

---

```
AIQPGGYCGYX-----CLSEQ ID NO:62
VSLMDRISQI-----GCSSEQ ID NO:63
VPAHLGPGAILDLK KY-----KYSEQ ID NO:64
FCYSTAGDGGSGE-----MYSSEQ ID NO:65
ELSGSCLEPLLE-----DYSEQ ID NO:66
DWKYWTGCAQTGGYF-----GSEQ ID NO:67
RLTEKGACDARWATLATRTFAANYSEQ ID NO:68
QKKDRTRWAEPREW-----NNSEQ ID NO:69
GRFSSPVGSTSLRES-SDY--KYSEQ ID NO:70
ADPSIYSLILEY-----KYSEQ ID NO:71
DSPCYMPTMPAPPIRDSFGW--DDSEQ ID NO:72
TSFYWYCTTAPY-----NVSSEQ ID NO:73
TEIEWYGCNLRITTF-----TRSEQ ID NO:74
NQLAGGWYLDENYWLVSQAY--ALSEQ ID NO:75
RLTEMGACDARWATLATRTFAANYSEQ ID NO:76
DGWTRKEGGIGLPWSVQCEGQYNYSEQ ID NO:77
DSYPCILL-----DVSSEQ ID NO:78
VEYFIADMCSE-----RYSSEQ ID NO:79
```

---

As stated above, the immunoglobulins of the invention are preferably devoid of the totality of their  $C_H1$  domain.

Such immunoglobulins comprise  $C_H2$  and  $C_H3$  domains in the C-terminal region with respect to the hinge region.

According to a particular embodiment of the invention the constant region of the immunoglobulins comprises  $C_H2$  and  $C_H3$  domains comprising an amino-acid sequence selected from the following for the  $C_H2$  domain:

---

```
APELLGGPTVFVFFPKPKDVLSTLTP SEQ ID NO:31
APELLGGGSPVFFVFFPKPKDVLSEGR SEQ ID NO:32
APELLGGGSPVFFVFFPKPKDVLSEGR SEQ ID NO:33
APELLGGGSPVFFVFFPKPKDVLSEGR SEQ ID NO:34
```

---

50 for the  $C_H3$  domain:

---

```
GOTREFQVYTLA SEQ ID NO:35
GOTREFQVYTHAPXRL SEQ ID NO:36
GOTREFQVYTHPPSRL SEQ ID NO:37
GOTREFQVYTHPPSRL SEQ ID NO:38
GOTREFQVYTHPPSRL SEQ ID NO:39
```

---

Interestingly the inventors have shown that the hinge region of the immunoglobulins of the invention can present variable lengths. When these immunoglobulins act as antibodies, the length of the hinge region will participate to the determination of the distance separating the antigen binding sites.

Preferably an immunoglobulin according to the invention is characterized in that its hinge region comprises from 10 to 50 amino-acids.

Particular sequences of hinge region of the immunoglobulins of the invention are the following.

GTNEVCRKCPKCP SEQ ID NO:37  
 EPKIPQPKPQPKPQPKPQPKPQPKPEPCTKCPKCP SEQ ID NO:38

The short hinge region corresponds to an IgG3 molecule and the long hinge sequence corresponds to an IgG2 molecule.

Isolated  $V_{HH}$  derived from heavy chain immunoglobulins or  $V_{HH}$  libraries corresponding to the heavy chain immunoglobulins can be distinguished from  $V_{HH}$  cloning of four-chain model immunoglobulins on the basis of sequence features characterizing heavy chain immunoglobulins.

The camel heavy-chain immunoglobulin  $V_{HH}$  region shows a number of differences with the  $V_{HH}$  regions derived from 4-chain immunoglobulins from all species examined. At the levels of the residues involved in the  $V_{HH}/V_L$  interactions, an important difference is noted at the level of position 45 (FW) which is practically always leucine in the 4-chain immunoglobulins (98%), the other amino acids at this position being proline (1%) or glutamine (1%).

In the camel heavy-chain immunoglobulin, in the sequences examined at present, leucine at position 45 is only found once. It could originate from a four-chain immunoglobulin. In the other cases, it is replaced by arginine, cysteine or glutamic acid residue. The presence of charged amino acids at this position should contribute to making the  $V_{HH}$  more soluble.

The replacement by camelid specific residues such as those of position 45 appears to be interesting for the construction of engineered  $V_{HH}$  regions derived from the  $V_{HH}$  repertoire of 4-chain immunoglobulins.

A second feature specific of the camelid  $V_{HH}$  domain is the frequent presence of a cysteine in the CDR<sub>3</sub> region associated with a cysteine in the CDR<sub>2</sub> position 31 or 33 or FW<sub>2</sub> region at position 45. The possibility of establishing a disulphide bond between the CDR<sub>2</sub> region and the rest of the variable domain would contribute to the stability and positioning of the binding site.

With the exception of a single pathogenic myeloma protein (DAW) such a disulphide bond has never been encountered in immunoglobulin V regions derived from 4 chain immunoglobulins.

The heavy-chain immunoglobulins of the invention have further the particular advantage of being not sticky. Accordingly these immunoglobulins being present in the serum, aggregate much less than isolated heavy chains of a four-chain immunoglobulins. The immunoglobulins of the invention are soluble to a concentration above 0.5 mg/ml, preferably above 1 mg/ml and more advantageously above 2 mg/ml.

These immunoglobulins further bear an extensive antigen binding repertoire and undergo affinity and specificity maturation in vivo. Accordingly they allow the isolation and the preparation of antibodies having defined specificity, regarding determined antigens.

Another interesting property of the immunoglobulins of the invention is that they can be modified and especially humanized. Especially it is possible to replace all or part of the constant region of these immunoglobulins by all or part of a constant region of a human antibody. For example the  $C_{H2}$  and/or  $C_{H3}$  domains of the immunoglobulin could be replaced by the  $C_{H2}$  and/or  $C_{H3}$  domains of the IgG  $\gamma 3$  human immunoglobulin.

In such humanized antibodies it is also possible to replace a part of the variable sequence, namely one or more of the

framework residues which do not intervene in the binding site by human framework residues, or by a part of a human antibody.

Conversely features (especially peptide fragments) of heavy-chain immunoglobulin  $V_{HH}$  regions, could be introduced into the  $V_H$  or  $V_L$  regions derived from four-chain immunoglobulins with for instance the aim of achieving greater solubility of the immunoglobulins.

The invention further relates to a fragment of an immunoglobulin which has been described hereabove and especially to a fragment selected from the following group:

a fragment corresponding to one heavy polypeptide chain of an immunoglobulin devoid of light chains,

fragments obtained by enzymatic digestion of the immunoglobulins of the invention, especially those obtained by partial digestion with pepsin leading to the Fc fragment (constant fragment) and leading to FV<sub>HH</sub> fragment (containing the antigen binding sites of the heavy chains) or its dimer F(V<sub>HH</sub>)<sub>2</sub>, or a fragment obtained by further digestion with papain of the Fc fragment, leading to the pFc fragment corresponding to the C-terminal part of the Fc fragment,

homologous fragments obtained with other proteolytic enzymes,

a fragment of at least 10 preferably 20 amino acids of the variable region of the immunoglobulin, or the complete variable region, especially a fragment corresponding to the isolated  $V_{HH}$  domains or to the  $V_{HH}$  dimers linked to the hinge disulphide,

a fragment corresponding to the hinge region of the immunoglobulin, or to at least 6 amino acids of this hinge region,

a fragment of the hinge region comprising a repeated sequence of Pro-X,

a fragment corresponding to at least 10 preferably 20 amino acids of the constant region or to the complete constant region of the immunoglobulin.

The invention also relates to a fragment comprising a repeated sequence, Pro-X which repeated sequence contains at least 3 repeats of Pro-X, X being any amino-acid and preferably Gln (glutamine), Lys (lysine) or Glu (acidic glutamic); a particular repeated fragment is composed of a 12-fold repeat of the sequence Pro-X.

Such a fragment can be advantageously used as a link between different types of molecules.

The amino-acids of the Pro-X sequence are chosen among any natural or non natural amino-acids.

The fragments can be obtained by enzymatic degradation of the immunoglobulins. They can also be obtained by expression in cells or organisms, of nucleotide sequence coding for the immunoglobulins, or they can be chemically synthesized.

The invention also relates to anti-idiotypes antibodies belonging to the heavy chain immunoglobulin classes. Such anti-idiotypes can be produced against human or animal idiotypes. A property of these anti-idiotypes is that they can be used as idiotype vaccines, in particular for vaccination against glycoproteins or glycolipids and where the carbohydrate determines the epitope.

The invention also relates to anti-idiotypes capable of recognizing idiotypes of heavy-chain immunoglobulins.

Such anti-idiotypic antibodies can be either syngeneic antibodies or allogeneic or xenogeneic antibodies.

The invention also concerns nucleotide sequences coding for all or part of a protein which amino-acid sequence comprises a peptide sequence selected from the following:



The examination of a large repertoire should prove to be particularly useful in the search for antibodies with catalytic activities.

The invention thus provides libraries which can be generated in a way which includes part of the hinge sequence, the identification is simple as the hinge is directly attached to the  $V_H$  domain.

These libraries can be obtained by cloning cDNA from lymphoid cells with or without prior PCR amplification. The PCR primers are located in the promoter, leader or framework sequences of the  $V_H$  from the 5' primer and in the hinge,  $CH_2$ ,  $CH_3$ , 3' untranslated region or polyA tail for the 3' primer. A size selection of amplified material allows the construction of a library limited to heavy chain immunoglobulins.

In a particular example, the following 3' primer in which a KpnI site has been constructed and which corresponds to amino-acids 313 to 319 (CGC CAT CAA GGT AAC AGT TGA)(SEQ ID NO:47) is used in conjunction with mouse  $V_H$  primers described by Sestry et al and containing a Xho site

```
AG CTC CAG CTG CTC GAG TCT GAG TCT GG SEQ ID NO:48
AG CTC CAG CTG CTC GAG TCT GAG TCT GG SEQ ID NO:49
AG GTC CAG CTG CTC GAG TCT GAG TCT GG SEQ ID NO:50
XhoI site
```

These primers yield a library of camelid heavy chain immunoglobulins comprising the  $V_H$  region (related to mouse or human subgroup III), the hinge and a section of  $CH_2$ .

In another example, the cDNA is polyadenylated at its 5' end and the mouse specific  $V_H$  primers are replaced by a poly T primer with an inbuilt XhoI site, at the level of nucleotide 12.

CTCGAGT<sub>12</sub>.

The same 3' primer with a KpnI site is used.

This method generates a library containing all subgroups of immunoglobulins.

Part of the interest in cloning a region encompassing the hinge- $CH_2$  link is that in both  $\gamma_2$  and  $\gamma_3$ , a SacI site is present immediately after the hinge. This site allows the grafting of the sequence coding for the  $V_H$  and the hinge onto the Fc region of other immunoglobulins, in particular the human IgG<sub>1</sub> and IgG<sub>3</sub> which have the same amino acid sequence at this site (Glu<sub>216</sub> Leu<sub>217</sub>).

As an example, the invention contemplates a cDNA library composed of nucleotide sequences coding for a heavy-chain immunoglobulin, such as obtained by performing the following steps:

- treating a sample containing lymphoid cells, especially peripheral, lymphocytes, spleen cells, lymph nodes or another Typhoid tissue from a healthy animal, especially selected among the Camelids, in order to separate the lymphoid cells,
- separating polyadenylated RNA from the other nucleic acids and components of the cells,
- reacting the obtained RNA with a reverse transcriptase in order to obtain the corresponding cDNA,
- contacting the cDNA of step c) with 5' primers corresponding to mouse  $V_H$  domain of four-chain immunoglobulins, which primer contains a determined restriction site, for example an XhoI site and with 3' primers corresponding to the N-terminal part of a  $C_H2$  domain containing a KpnI site,

e) amplifying the DNA,

f) cloning the amplified sequence in a vector, especially in a bluescript vector,

g) recovering the clones hybridizing with a probe corresponding to the sequence coding for a constant domain from an isolated heavy-chain immunoglobulin.

This cloning gives rise to clones containing DNA sequences including the sequence coding for the hinge. It thus permits the characterization of the subclass of the immunoglobulin and the SacI site useful for grafting the FV<sub>H</sub> to the Fc region.

The recovery of the sequences coding for the heavy-chain immunoglobulins can also be achieved by the selection of clones containing DNA sequences having a size compatible with the lack of the  $C_H1$  domain.

It is possible according to another embodiment of the invention, to add the following steps between steps c) and d) of the above process:

in the presence of a DNA polymerase and of deoxyribonucleotide triphosphates, contacting said cDNA with oligonucleotide degenerated primers, which sequences are capable of coding for the hinge region and N-terminal  $V_H$  domain of an immunoglobulin, the primers being capable of hybridizing with the cDNA and capable of initiating the extension of a DNA sequence complementary to the cDNA used as template,

recovering the amplified DNA.

The clones can be expressed in several types of expression vectors. As an example using a commercially available vector Immuno PBS (Huse et al: Science (1989) 246, 1275), clones produced in Bluescript® according to the above described procedure, are recovered by PCR using the same XhoI containing 5' primer and a new 3' primer, corresponding to residues 113-103 in the framework of the immunoglobulins, in which an Spe site has been constructed: TC TTA ACT AGT GAG GAG ACG GTG ACC TG (SEQ ID NO:51). This procedure allows the cloning of the  $V_H$  in the Xho/Spe site of the Immuno PBS vector. However, the 3' end of the gene is not in phase with the identification "tag" and the stop codon of the vector. To achieve this, the construct is cut with Spe and the 4 base overhangs are filled in, using the Klenow fragment after which the vector is religated. A further refinement consists in replacing the marker ("tag") with a poly histidine so that a metal purification of the cloned  $V_H$  can be performed. To achieve this a Spe/EcoRI double stranded oligonucleotide coding for 6 histidines and a termination codon is first constructed by synthesis of both strands followed by heating and annealing:

```
CTA CTC CAG CAC CAC CAT CAC CAT CAC TAA TAG SEQ ID NO:52
AC GTG GTG GTA GTG GTA GTG ATT ATC TTA SEQ ID NO:53
```

The vector containing the insert is then digested with SpeI and EcoRI to remove the resident "tag" sequence which can be replaced by the poly-His/termination sequence. The produced  $V_H$  can equally be detected by using antibodies raised against the dromedary  $V_H$  regions. Under laboratory conditions,  $V_H$  regions are produced in the Immuno PBS vector in mg amounts per liter.

The invention also relates to a DNA library composed of nucleotide sequences coding for a heavy-chain immunoglobulin, such as obtained from cells with rearranged immunoglobulin genes.

In a preferred embodiment of the invention, the library is prepared from cells from an animal previously immunized



against a determined antigen. This allows the selection of antibodies having a preselected specificity for the antigen used for immunization.

In another embodiment of the invention, the amplification of the cDNA is not performed prior to the cloning of the cDNA.

The heavy-chain of the four-chain immunoglobulins remains sequestered in the cell by a chaperon protein (BIP) until it has combined with a light chain. The binding site for the chaperon protein is the  $C_H1$  domain. As this domain is absent from the heavy chain immunoglobulins, their secretion is independent of the presence of the BIP protein or of the light chain. Moreover the inventors have shown that the obtained immunoglobulins are not sticky and accordingly will not abnormally aggregate.

The invention also relates to a process for the preparation of a monoclonal antibody directed against a determined antigen, the antigen binding site of the antibody consisting of heavy polypeptide chains and which antibody is further devoid of light polypeptide chains, which process comprises:

immortalizing lymphocytes, obtained for example from the peripheral blood of Camelids previously immunized with a determined antigen, with an immortal cell and preferably with myeloma cells, in order to form a hybridoma,

culturing the immortalized cells (hybridoma) formed and recovering the cells producing the antibodies having the desired specificity.

The preparation of antibodies can also be performed without a previous immunization of Camelids.

According to another process for the preparation of antibodies, the recourse to the technique of the hybridoma cell is not required.

According to such process, antibodies are prepared in vitro and they can be obtained by a process comprising the steps of:

cloning into vectors, especially into phages and more particularly filamentous bacteriophages, DNA or cDNA sequences obtained from lymphocytes especially PBLs of Camelids previously immunized with determined antigens,

transforming prokaryotic cells with the above vectors in conditions allowing the production of the antibodies, selecting the antibodies for their heavy-chain structure and further by subjecting them to antigen-affinity selection,

recovering the antibodies having the desired specificity.

In another embodiment of the invention the cloning is performed in vectors, especially into plasmids coding for bacterial membrane proteins. Prokaryotic cells are then transformed with the above vectors in conditions allowing the expression of antibodies in their membrane.

The positive cells are further selected by antigen affinity selection.

The heavy chain antibodies which do not contain the  $C_H1$  domain present a distinct advantage in this respect. Indeed, the  $C_H1$  domain binds to BIP type chaperone proteins present within eukaryotic vectors and the heavy chains are not transported out of the endocytoplasmic reticulum unless light chains are present. This means that in eukaryotic cells, efficient cloning of 4-chain immunoglobulins in non mammalian cells such as yeast cells can depend on the properties of the resident BIP type chaperone and can hence be very difficult to achieve. In this respect the heavy chain antibodies of the invention which lack the  $C_H1$  domain present a distinctive advantage.

In a preferred embodiment of the invention the cloning can be performed in yeast either for the production of antibodies or for the modification of the metabolism of the yeast. As example, Yep 52 vector can be used. This vector has the origin of replication (ORI)  $2\mu$  of the yeast together with a selection marker *Leu 2*.

The cloned gene is under the control of *galI* promoter and accordingly is inducible by galactose. Moreover, the expression can be repressed by glucose which allows the obtention of very high concentration of cells before the induction.

The cloning between BamHI and SalI sites using the same strategy of production of genes by PCR as the one described above, allows the cloning of camelid immunoglobulin genes in *E. coli*. As example of metabolic modulation which can be obtained by antibodies and proposed for the yeast, one can site the cloning of antibodies directed against cyclins, that is proteins involved in the regulation of the cellular cycle of the yeast (TIBS 16 430 J. D. McKinney, N. Heintz 1991). Another example is the introduction by genetic engineering of an antibody directed against CD<sub>28</sub> which antibody would be inducible (for instance by gal), within the genome of the yeast. The CD<sub>28</sub> is involved at the level of the initiation of cell division, and therefore the expression of antibodies against this molecule would allow an efficient control of multiplication of the cells and the optimization of methods for the production in bioreactors or by means of immobilized cells.

In yet another embodiment of the invention, the cloning vector is a plasmid or a eukaryotic virus vector and the cells to be transformed are eukaryotic cells, especially yeast cells, mammalian cells for example CHO cells or simian cells such as Vero cells, insect cells, plant cells, or protozoan cells.

For more details concerning the procedure to be applied in such a case, reference is made to the publication of Marks et al, J. Mol. Biol. 1991, 222:581-597.

Furthermore, starting from the immunoglobulins of the invention, or from fragments thereof, new immunoglobulins or derivatives can be prepared.

Accordingly immunoglobulins replying to the above given definitions can be prepared against determined antigens. Especially the invention provides monoclonal or polyclonal antibodies devoid of light polypeptide chains or antisera containing such antibodies and directed against determined antigens and for example against antigens of pathological agents such as bacteria, viruses or parasites. As example of antigens or antigenic determinants against which antibodies could be prepared, one can cite the envelope glycoproteins of viruses or peptides thereof, such as the external envelope glycoprotein of a HIV virus, the surface antigen of the hepatitis B virus.

Immunoglobulins of the invention can also be directed against a protein, hapten, carbohydrate or nucleic acid.

Particular antibodies according to the invention are directed against the galactosyl-1-3-galactose epitope.

The immunoglobulins of the invention allow further the preparation of combined products such as the combination of the heavy-chain immunoglobulin or a fragment thereof with a toxin, an enzyme, a drug, a hormone.

As example one can prepare the combination of a heavy-chain immunoglobulin bearing an antigen binding site recognizing a myeloma immunoglobulin epitope with the abrin or mistletoe lectin toxin. Such a construct would have its uses in patient specific therapy.

Another advantageous combination is that one can prepare between a heavy-chain immunoglobulin recognizing an insect gut antigen with a toxin specific for insects such as

the toxins of the different serotypes of *Bacillus thuringiensis* or *Bacillus sphaericus*. Such a construct cloned into plants can be used to increase the specificity or the host range of existing bacterial toxins.

The invention also proposes antibodies having different specificities on each heavy polypeptide chains. These multifunctional, especially bifunctional antibodies could be prepared by combining two heavy chains of immunoglobulins of the invention or one heavy chain of an immunoglobulin of the invention with a fragment of a four-chain model immunoglobulin.

The invention also provides hetero-specific antibodies which can be used for the targeting of drugs or any biological substance like hormones. In particular they can be used to selectively target hormones or cytokines to a limited category of cells. Examples are a combination of a murine or human antibody raised against interleukin 2 (IL<sub>2</sub>) and a heavy-chain antibody raised against CD<sub>3</sub> cells. This could be used to reactivate CD<sub>3</sub> cells which have lost their IL<sub>2</sub> receptor.

The heavy-chain immunoglobulins of the invention can also be used for the preparation of hetero-specific antibodies. These can be achieved either according to the above described method by reduction of the bridges between the different chains and reoxidation, according to the usual techniques, of two antibodies having different specificities, but it can also be achieved by serial cloning of two antibodies for instance in the Immuno pBS vector.

In such a case, a first gene corresponding to the  $V_{H1}$  domain comprised between Xho site and a Spe site is prepared as described above. A second gene is then prepared through an analogous way by using as 5' extremity a primer containing a Spe site, and as 3' extremity a primer containing a termination codon and an EcoRI site. The vector is then digested with EcoRI and XhoI and further both  $V_{H1}$  genes are digested respectively by Xho/Spe and by Spe/EcoRI.

After ligation, both immunoglobulin genes are serially cloned. The spacing between both genes can be increased by the introduction of addition codons within the 5' Spel primer.

In a particular embodiment of the invention, the hinge region of IgG2 immunoglobulins according to the invention is semi-rigid and is thus appropriate for coupling proteins. In such an application proteins or peptides can be linked to various substances, especially to ligands through the hinge region used as spacer. Advantageously the fragment comprises at least 6 amino acids.

According to the invention it is interesting to use a sequence comprising a repeated sequence Pro-X, X being any amino-acid and preferably Gln, Lys or Glu, especially a fragment composed of at least a 3-fold repeat and preferably of a 12-fold repeat, for coupling proteins to ligand, or for assembling different protein domains.

The hinge region or a fragment thereof can also be used for coupling proteins to ligands or for assembling different protein domains.

Usual techniques for the coupling are appropriate and especially reference may be made to the technique of protein engineering by assembling cloned sequences.

The antibodies according to this invention could be used as reagents for the diagnosis in vitro or by imaging techniques. The immunoglobulins of the invention could be labelled with radio-isotopes, chemical or enzymatic markers or chemiluminescent markers.

As example and especially in the case of detection or observation with the immunoglobulins by imaging techniques, a label like technetium, especially technetium 99 is advantageous. This label can be used for direct labelling

by a coupling procedure with the immunoglobulins or fragments thereof or for indirect labelling after a step of preparation of a complex with the technetium.

Other interesting radioactive labels are for instance indium and especially indium 111, or iodine, especially I<sup>131</sup>, I<sup>125</sup> and I<sup>123</sup>.

For the description of these techniques reference is made to the FR patent application published under number 2649488.

In these applications the small size of the  $V_{H1}$  fragment is a definitive advantage for penetration into tissue.

The invention also concerns monoclonal antibodies reacting with anti-idiotypes of the above-described antibodies.

The invention also concerns cells or organisms in which heavy-chain immunoglobulins have been cloned. Such cells or organisms can be used for the purpose of producing heavy-chain immunoglobulins having a desired proselected specificity, or corresponding to a particular repertoire. They can also be produced for the purpose of modifying the metabolism of the cell which expresses them. In the case of modification of the metabolism of cells transformed with the sequences coding for heavy-chain immunoglobulins, these produced heavy-chain immunoglobulins are used like antisense DNA. Antisense DNA is usually involved in blocking the expression of certain genes such as for instance the variable surface antigen of trypanosomes or other pathogens. Likewise, the production or the activity of certain proteins or enzymes could be inhibited by expressing antibodies against this protein or enzyme within the same cell.

The invention also relates to a modified 4-chain immunoglobulin or fragments thereof, the  $V_H$  regions of which has been partially replaced by specific sequences or amino acids of heavy chain immunoglobulins, especially by sequences of the  $V_{H1}$  domain. A particular modified  $V_H$  domain of a four-chain immunoglobulin, is characterized in that the leucine, proline or glutamine in position 45 of the  $V_H$  regions has been replaced by other amino acids and preferably by arginine, glutamic acid or cysteine.

A further modified  $V_H$  or  $V_L$  domain of a four-chain immunoglobulin, is characterized by linking of CDR loops together or to FW regions by the introduction of paired cysteines, the CDR region being selected between the CDR<sub>1</sub> and the CDR<sub>2</sub>, the FW region being the FW<sub>2</sub> region, and especially in which one of the cysteines introduced is in position 31, 33 of the CDR<sub>1</sub> or 45 of FW<sub>2</sub> and the other in CDR<sub>2</sub>.

Especially the introduction of paired cysteines is such that the CDR<sub>1</sub> loop is linked to the FW<sub>2</sub> or CDR1 domain and more especially the cysteine of the CDR3 of the  $V_H$  is linked to a cysteine in position 31, 33 of CDR<sub>1</sub> or in position 45 of FW<sub>2</sub>.

In another embodiment of the invention, plant cells can be modified by the heavy-chain immunoglobulins according to the invention, in order that they acquire new properties or increased properties.

The heavy-chain immunoglobulins of the invention can be used for gene therapy of cancer for instance by using antibodies directed against proteins present on the tumor cells.

In such a case, the expression of one or two  $V_{H1}$  genes can be obtained by using vectors derived from parvo or adeno viruses. The parvo viruses are characterized by the fact that they are devoid of pathogenicity or almost not pathogenic for normal human cells and by the fact that they are capable of easily multiplying in cancer cells (Russell S. J. 1990, Immunol. Today II, 196-200).

The heavy-chain immunoglobulins are for instance cloned within HindIII/XbaI sites of the infectious plasmid of

the murine MVM virus (pMM984). (Merchinsky et al, 1983, J. Virol, 47, 227-232) and then placed under the control of the MVM38 promoter.

The gene of the  $V_{HII}$  domain is amplified by PCR by using a 5' primer containing an initiation codon and a HindIII site, the 3' primer containing a termination codon and a XbaI site.

This construct is then inserted between positions 2650 (HindIII) and 4067 (XbaI) within the plasmid.

The efficiency of the cloning can be checked by transfection. The vector containing the antibody is then introduced in permissive cells (NB-E) by transfection.

The cells are recovered after two days and the presence of  $V_{HII}$  regions is determined with an ELISA assay by using rabbit antiserum reacting with the  $V_{HII}$  part.

The invention further allows the preparation of catalytic antibodies through different ways. The production of antibodies directed against components mimicking activated states of substrates (as example vanadate as component mimicking the activated state of phosphate in order to produce their phosphoesterase activities, phosphonate as compound mimicking the peptidic binding in order to produce proteases) permits to obtain antibodies having a catalytic function. Another way to obtain such antibodies consists in performing a random mutagenesis in clones of antibodies for example by PCR, in introducing abnormal bases during the amplification of clones. These amplified fragments obtained by PCR are then introduced within an appropriate vector for cloning. Their expression at the surface of the bacteria permits the detection by the substrate of clones having the enzymatic activity. These two approaches can of course be combined. Finally, on the basis of the data available on the structure, for example the data obtained by XRay crystallography or NMR, the modifications can be directed. These modifications can be performed by usual techniques of genetic engineering or by complete synthesis. One advantage of the  $V_{HII}$  of the heavy chain immunoglobulins of the invention is the fact that they are sufficiently soluble.

The heavy chain immunoglobulins of the invention can further be produced in plant cells, especially in transgenic plants. As example the heavy chain immunoglobulins can be produced in plants using the pMon530 plasmid (Roger et al. Meth Enzym 153 1566 1987) constitutive plant expression vector as has been described for classical four chain antibodies (Hiat et al. Nature 342 76-78, 1989) once again using the appropriate PCR primers as described above, to generate a DNA fragment in the right phase.

Other advantages and characteristics of the invention will become apparent in the examples and figures which follow.

## FIGURES

FIGS. 1(A), (B), and (C) Characterisation and purification of camel IgG by affinity chromatography on Protein A and Protein G sepharose (Pharmacia)

FIG. (A) shows, after reduction, the SDS-PAGE protein profile of the adsorbed and not adsorbed fractions of *Camelus dromedarius* serum. The fraction adsorbed on Protein A and eluted with NaCl 0.15M acetic acid 0.58% show upon reduction (lane c) three heavy chain components of respectively 50, 46 and 43 Kd and light chain (rabbit IgG in lane a). The fractions adsorbed on a Protein G Sepharose (Pharmacia) derivative which has been engineered to delete the albumin binding region (lane c) and eluted with 0.1M gly HCl pH 2.7 lacks the 46 Kd heavy chain which is recovered in the non adsorbed fraction (lane f). None of these components are present in the fraction non adsorbed on Protein A (lane d), lane b contains the molecular weight markers.

FIGS. (B) and (C) By differential elution, immunoglobulin fractions containing the 50 and 43 Kd heavy chain can be separated. 5 ml of *C. dromedarius* serum is adsorbed onto a 5 ml Protein G sepharose column and the column is extensively washed with 20 mM phosphate buffer, pH 7.0. Upon elution with pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid) a 100 Kd component is eluted which upon reduction yields a 43 Kd heavy chain, (lane 1). After column eluant absorbance has fallen to background level a second immunoglobulin component of 170 Kd can be eluted with pH 2.7 buffer (0.1M glycine HCl). This fraction upon reduction yields a 50 Kd heavy chain and a broad light chain band (lane 2). The fraction non adsorbed on Protein G is then brought on a 5 ml Protein A Sepharose column. After washing and elution with pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid) a third immunoglobulin of 100 Kd is obtained which consists solely of 46 Kd heavy chains (lane 3).

FIGS. 2(A) and 2(B) Immunoglobulins of *Camelus bactrianus*, *Lama vicugna*, *Lama glama* and *Lama pacos* to Protein A (A lanes) and to Protein G (G lanes) analyzed on SDS-PAGE before FIG. (A) and after reduction FIG. (B)

10  $\mu$ l of serum obtained from the different species were added to Eppendorf® tubes containing 10 mg of Protein A or Protein G sepharose suspended in 400  $\mu$ l of pH 8.3 immunoprecipitation buffer (NaCl 0.2M, Tris 0.01M; EDTA 0.01M, Triton X100 1%, ovalbumin 0.1%). The tubes were slowly rotated for 2 hours at 4° C. After centrifugation the pellets were washed 3 times in buffer and once in buffer in which the Triton and ovalbumin had been omitted. The pellets were then resuspended in the SDS-PAGE sample solution 70  $\mu$ l per pellet with or without dithiothreitol as reductant. After boiling for 3 min at 100° C., the tubes were centrifuged and the supernatants analysed. In all species examined the unreduced fractions FIG. (A) contain in addition to molecules of approximately 170 Kd also smaller major components of approximately 100 Kd. In the reduced sample FIG. (B) the constituent heavy and light chains are detected. In all species a heavy chain component (marked by an asterisk \*) is present in the material eluted from the Protein A but absent in the material eluted from the Protein G.

FIGS. 3(A)-(C): IgG<sub>1</sub>, IgG<sub>2</sub> and IgG<sub>3</sub> were prepared from serum obtained from healthy or *Trypanosoma evansi* infected *Camelus dromedarius* (CAIT tier 1/160) (3) and analysed by radioimmunoprecipitation or Western Blotting for anti trypanosome activity

FIG. (A) <sup>35</sup>S methionine labelled *Trypanosoma evansi* antigens lysate (500,000 counts) was added to Eppendorf tubes containing 10  $\mu$ l of serum or, 20  $\mu$ g of IgG<sub>1</sub>, IgG<sub>2</sub> or IgG<sub>3</sub> in 200  $\mu$ l of pH 8.3 immunoprecipitation buffer containing 0.1M TLCK as proteinase inhibitor and slowly rotated at 4° C. during one hour. The tubes were then supplemented with 10 mg of Protein A Sepharose suspended in 200  $\mu$ l of the same pH 8.3 buffer and incubated at 4° C. for an additional hour.

After washing and centrifugation at 15000 rpm for 12 s, each pellet was resuspended in 75  $\mu$ l SDS-PAGE sample solution containing DTT and heated for 3 min. at 100° C. After centrifugation in a Eppendorf minifuge at 15000 rpm for 30 s, 5  $\mu$ l of the supernatant was saved for radioactivity determination and the remainder analysed by SDS-PAGE and fluorography. The counts/5  $\mu$ l sample are inscribed on for each line.

FIGS. (B) and (C) 20  $\mu$ g of IgG<sub>1</sub>, IgG<sub>2</sub> and IgG<sub>3</sub> from healthy and trypanosome infected animals were separated by SDS-PAGE without prior reduction or heating. The sepa-

rated samples were then electro transferred to a nitrocellulose membrane, one part of the membrane was stained with Ponceau Red to localise the protein material and the remainder incubated with 1% ovalbumin in TST buffer (Tris 10 mM, NaCl 150 mM, Tween 0.05%) to block protein binding sites.

After blocking, the membrane was extensively washed with TST buffer and incubated for 2 hours with  $^{35}$ S-labelled trypanosome antigen. After extensive washing, the membrane was dried and analysed by autoradiography. To avoid background and unspecific binding, the labelled trypanosome lysate was filtered through a 45  $\mu$  millipore filter and incubated with healthy camel immunoglobulin and ovalbumin adsorbed on a nitrocellulose membrane.

FIGS. 4(A) and (B): Purified IgG3 of the camel, by affinity chromatography on Protein A Sepharose are partially digested with papain and separated on Protein A sepharose.

14 mg of purified IgG3 were dissolved in 0.1M phosphate buffer pH 7.0 containing 2mM EDTA. They were digested by 1 hour incubation at 37° C. with mercurypapain (1% enzyme to protein ratio) activated by 5.10<sup>-3</sup>M cysteine. The digestion was blocked by the addition of excess iodoacetamide (4.10<sup>-2</sup>M) (13). After centrifugation of the digest in an endorf centrifuge for 5 min at 15000 rpm, the papain fragments were separated on a protein A Sepharose column into binding (B) and non binding (NB) fractions. The binding fraction was eluted from the column with 0.1M glycine HCl buffer pH 1.7.

FIG. 5: Schematic presentation of a model for IgG3 molecules devoid of light chains.

FIG. 6: Schematic representation of immunoglobulins having heavy polypeptide chains and devoid of light chains, regarding conventional four-chain model immunoglobulin. Representation of a hinge region.

FIG. 7: Alignment of 17 V<sub>H</sub> DNA sequences of Camel heavy chain immunoglobulins SEI ID NOS:92-108

FIGS. 8(A) and (B): Expression and purification of the camel V<sub>H</sub>21 protein from *E. coli*

# I HEAVY CHAIN ANTIBODIES IN CAMELIDS

When *Camelus dromedarius* serum is adsorbed on Protein G sepharose, an appreciable amount (25-35%) of immunoglobulins (Ig) remains in solution which can then be recovered by affinity chromatography on Protein A sepharose (FIG. 1A). The fraction adsorbed on Protein G can be differentially eluted into a tightly bound fraction (25%) consisting of molecules of an unreduced apparent molecular weight (MW) of 170 Kd and a more weakly bound fraction (30-45%) having an apparent molecular weight of 100 Kd (FIG. 1B). The 170 Kd component when reduced yields 50 Kd heavy chains and large 30 Kd light chains. The 100 Kd fraction is totally devoid of light chains and appears to be solely composed of heavy chains which after reduction have on apparent MW of 43 Kd (FIG. 1C). The fraction which does not bind to Protein G can be affinity purified and eluted from a Protein A column as a second 100 Kd component which after reduction appears to be composed solely of 46 Kd heavy chains.

The heavy chain immunoglobulins devoid of light chains total up to 75% of the molecules binding to Protein A.

As all three immunoglobulins bind to Protein A we refer to them as IgG: namely IgG<sub>1</sub> (light chain and heavy chain  $\gamma$ 1 (50 Kd) binding to Protein G, IgG2 (heavy chain  $\gamma$ 2 (46 Kd) non binding to Protein G and IgG3 (heavy chain  $\gamma$ 3 (43 Kd) binding to Protein G. There is a possibility that these three sub(classes) can be further subdivided.

A comparative study of old world camelids (*Camelus bactrianus* and *Camelus dromedarius*) and new world camelids (*Lama pacos*, *Lama glama*, *Lama vicugna*) showed that heavy chain immunoglobulins are found in all species examined, albeit with minor differences in apparent molecular weight and proportion. The new world camelids differs from the old world camelids in having a larger IgG<sub>2</sub> molecule (heavy chain immunoglobulin binding to Protein G) in which the constituent heavy chains have an apparent molecular weight of 47 Kd (FIGS. 2A and B).

The abundance of the heavy chain immunoglobulins in the serum of camelids raises the question of what their role is in the immune response and in particular whether they bear antigen binding specificity and if so how extensive is the repertoire. This question could be answered by examining the immunoglobulins from *Trypanosoma evansi* infected camels (*Camelus dromedarius*).

For this purpose, the corresponding fractions of IgG<sub>1</sub>, IgG<sub>2</sub>, IgG<sub>3</sub> were prepared from the serum of a healthy camel, and from the serum of camels with a high antitrypanosome titer, measured by the Card Agglutination Test (3). In radioimmunoprecipitation, IgG<sub>1</sub>, IgG<sub>2</sub> and IgG<sub>3</sub> derived from infected camel indicating extensive repertoire heterogeneity and complexity (FIG. 3A) were shown to bind a large number of antigens present in a  $^{35}$ S methionine labelled trypanosome lysate.

In blotting experiments  $^{35}$ S methionine labelled trypanosome lysate binds to SDS PAGE separated IgG<sub>1</sub>, IgG<sub>2</sub> and IgG<sub>3</sub> obtained from infected animals (FIG. 3B).

This leads us to conclude that the camelid heavy chain IgG<sub>2</sub> and IgG<sub>3</sub> are bona fide antigen binding antibodies.

An immunological paradigm states that an extensive antibody repertoire is generated by the combination of the light and heavy chain variable V region repertoires (6). The heavy chain immunoglobulins of the camel seem to contradict this paradigm.

Immunoglobulins are characterized by a complex I.E.F. (isoelectric focussing) pattern reflecting their extreme heterogeneity. To determine whether the two heavy chains constituting the IgG<sub>2</sub> and IgG<sub>3</sub> are identical or not, the isoelectric focussing (I.E.F) pattern were observed before and after chain separation by reduction and alkylation using iodoacetamide as alkylating agent.

As this alkylating agent does not introduce additional changes in the molecule, the monomers resulting from the reduction and alkylation of a heavy chain homodimer will have practically the same isoelectric point as the dimer, whereas if they are derived from a heavy chain heterodimer, the monomers will in most cases differ sufficiently in isoelectric point to generate a different pattern in I.E.F.

Upon reduction, and alkylation by iodoacetamide the observed pattern is not modified for the *Camelus dromedarius* IgG<sub>2</sub> and IgG<sub>3</sub>, indicating that these molecules are each composed of two identical heavy chains which migrate to the same position as the unreduced molecule they originated from.

In contrast, the I.E.F. pattern of IgG<sub>1</sub> is completely modified after reduction as the isoelectric point of each molecule is determined by the combination of the isoelectric points of the light and heavy chains which after separation will each migrate to a different position.

These findings indicate that the heavy chains alone can generate an extensive repertoire and question the contribution of the light chain to the useful antibody repertoire. If this necessity be negated, what other role does the light chain play.

Normally, isolated heavy chain from mammalian immunoglobulins tend to aggregate considerably but are only solubilized by light chains (8, 9) which bind to the  $C_{H1}$  domain of the heavy chain.

In humans and in mice a number of spontaneous or induced myelomas produce a pathological immunoglobulin solely composed of heavy chains (heavy chain disease). These myeloma protein heavy chains carry deletions in the  $C_{H1}$  and  $V_{HH}$  domains (10). The reason why full length heavy chains do not give rise to secreted heavy chain in such pathological immunoglobulins seems to stem from the fact that the synthesis of Ig involves a chaperoning protein, the immunoglobulin heavy chain binding protein or BIP (11), which normally is replaced by the light chain (12). It is possible that the primordial role of the light chain in the four-chain model immunoglobulins is that of a committed heavy chain chaperon and that the emergence of light chain repertoires has just been an evolutionary bonus.

The camelid  $\gamma 2$  and  $\gamma 3$  chains are considerably shorter than the normal mammalian  $\gamma$  chain. This would suggest that deletions have occurred in the  $C_{H1}$  domain. Differences in sizes of the  $\gamma 2$  and  $\gamma 3$  immunoglobulins of old and new world camelids suggests that deletions occurred in several evolutionary steps especially in the  $C_{H1}$  domain.

#### II THE HEAVY CHAIN IMMUNOGLOBULINS OF THE CAMELIDS LACK THE $C_{H1}$ DOMAIN

The strategy followed for investigating the heavy chain immunoglobulin primary structure is a combination of protein and cDNA sequencing: the protein sequencing is necessary to identify sequence stretches characteristic of each immunoglobulin. The N-terminal of the immunoglobulin being derived from the heavy chain variable region repertoire only yields information on the  $V_{HH}$  subgroups (variable region of the heavy chain) and cannot be used for class or subclass identification. This means that sequence data had to be obtained from internal enzymatic or chemical cleavage sites.

A combination of papain digestion and Protein A affinity chromatography allowed the separation of various fragments yielding information on the general structure of IgG3.

The IgG3 of the camel (*Camelus dromedarius*) purified by affinity chromatography on Protein A Sepharose were partially digested with papain and the digest was separated on Protein A Sepharose into binding and non binding fractions. These fractions were analysed by SDS PAGE under reducing and non reducing conditions (FIG. 4A and B).

The bound fraction contained two components, one of 28 Kd and one of 14.4 Kd, in addition to uncleaved or partially cleaved material. They were well separated by gel electrophoresis (from preparative 19% SDS-PAGE gels) under non reducing conditions and were further purified by electroelution (in 50 mM ammonium bicarbonate, 0.1% (w/v) SDS using a BioRad electro-eluter). After lyophilization of these electroeluted fractions, the remaining SDS was eliminated by precipitating the protein by the addition of 90% ethanol, mixing and incubating the mixture overnight at  $-20^{\circ}\text{C}$ . (14). The precipitated protein was collected in a pellet by centrifuging (15000 rpm, 5 min) and was used for protein sequencing. N-terminal sequencing was performed using the automated Edman chemistry of an Applied Biosystem 477A pulsed liquid protein sequencer. Amino acids were identified as their phenylthiohydantoin (PTHT) derivatives using an Applied Biosystem 120 PTH analyser. All chemical and reagents were purchased from Applied Biosystems. Analysis of the chromatographic data was performed using Applied

Biosystems software version 1.61. In every case the computer aided sequence analysis was confirmed by direct inspection of the chromatograms from the PTH analyser. Samples for protein sequencing were dissolved in either 50% (v/v) trifluoroacetic acid (TFA) (28 Kd fragment) or 100% TFA (14 Kd fragment). Samples of dissolved protein equivalent to 2000 pmol (28 Kd fragment) or 500 pmol (14 Kd fragment) were applied to TFA-treated glass fibre discs. The glass fibre discs were coated with BioBrene (3mg) and precycled once before use.

N-terminal sequencing of the 28 Kd fragment yields a sequence homologous to the N-terminal part of  $\gamma C_{H2}$  domain and hence to the N-terminal end of the Fc fragment. The N-terminal sequence of the 14.4 Kd fragment corresponds to the last lysine of a  $\gamma C_{H2}$  and the N-terminal end of a  $\gamma C_{H3}$  domain (Table 1). The molecular weight (MW) of the papain fragments and the identification of their N-terminal sequences led us to conclude that the  $C_{H2}$  and  $C_{H3}$  domains of the  $\gamma 3$  heavy chains are normal in size and that the deletion must occur either in the  $C_{H1}$  or in the  $V_{HH}$  domain to generate the shortened  $\gamma 3$  chain. The fractions which do not bind to Protein A Sepharose contain two bands of 34 and 17 Kd which are more diffuse is SDS PAGE indicating that they originate from the variable N-terminal part of the molecule (FIGS. 4A and B).

Upon reduction, a single diffuse band of 17 Kd is found indicating that the 34 Kd is a disulfide bonded dimer of the 17 Kd component. The 34 Kd fragment apparently contains the hinge and the N-terminal domain  $V_{HH}$ .

The protein sequence data can be used to construct degenerate oligonucleotide primers allowing PCR amplification of cDNA or genomic DNA.

It has been shown that the cells from camel spleen imprint cells reacted with rabbit and anti camel immunoglobulin sera and that the spleen was hence a site of synthesis of at least one immunoglobulin class. cDNA was therefore synthesised from camel spleen mRNA. The conditions for the isolation of RNA were the following: total RNA was isolated from the dromedary spleen by the guanidium isothiocyanate method (15). mRNA was purified with oligo T-paramagnetic beads.

cDNA synthesis is obtained using 1  $\mu\text{g}$  mRNA template, an oligodT primer and reverse transcriptase (BOEHRINGER MANN). Second strand cDNA is obtained using RNase H and *E. coli* DNA polymerase I according to the condition given by the supplier.

Relevant sequences were amplified by PCR: 5 ng of cDNA was amplified by PCR in a 100  $\mu\text{l}$  reaction mixture (10 mM Tris-HCl pH 8.3, 50  $\mu\text{M}$  KCl, 15 mM  $\text{MgCl}_2$ , 0.01% (v/v) gelatine, 200  $\mu\text{M}$  of each dNTP and 25 pmol of each primer) overlaid with mineral oil (Sigma). Degenerate primers containing EcoRI and KpnI sites and further cloned into pUC 18. After a round of denaturing and annealing ( $94^{\circ}\text{C}$ . for 5 min and  $54^{\circ}\text{C}$ . for 5 min), 2 units of Taq DNA polymerase were added to the reaction mixture before subjecting it to 35 cycles of amplification: 1 min at  $94^{\circ}\text{C}$ . (denature) 1 min at  $54^{\circ}\text{C}$ . (anneal), 2 min at  $72^{\circ}\text{C}$ . (elongate). To amplify DNA sequences between  $V_{HH}$  and  $C_{H2}$  domains, (#72 clones), the PCR was performed in the same conditions with the exception that the annealing temperature was increased to  $60^{\circ}\text{C}$ .

One clone examined (#56/36) had a sequence corresponding to the N-terminal part of a  $C_{H2}$  domain identical to the sequence of the 28 Kd fragment. The availability of this sequence data allowed the construction of an exact 3' primer and the cloning of the region between the N-terminal end of the  $V_{HH}$  and the  $C_{H2}$  domain.

5' primers corresponding to the mouse  $V_{H\mu}$  (16) and containing a XhoI restriction site were used in conjunction with the 3' primer in which a KpnI site had been inserted and the amplified sequences were cloned into pBluescript®. Clone #56/36 which displayed two internal HaeIII sites was digested with this enzyme to produce a probe to identify PCR positive clones.

After amplification the PCR products were checked on a 1.2% (w/v) agarose gel. Cleaning up of the PCR products included a phenol-chloroform extract followed by further purification by HPLC (GEN-PAC FAX column, Waters) and finally by using the MERMAID or GENECLAN II kit, BIO 101, Inc.) as appropriate. After these purification steps, the amplified cDNA was then digested with EcoRI and KpnI for series #56 clones and with XhoI and KpnI for series #72 clones. A final phenol-chloroform extraction preceded the ligation into pUC 18 (series #56 clones) or into pBluescript® (series #72 clones).

All the clones obtained were smaller than the 860 base pairs to be expected if they possessed a complete  $V_{H\mu}$  and  $C_{H\mu}$  region. Partial sequence data corresponding to the N-terminal of the  $V_{H\mu}$  region reveals that out of 20 clones, 3 were identical and possibly not independent. The sequences obtained resemble the human subgroup III and the murine subgroups IIIa and IIIb (Table 2).

Clones corresponding to two different sets of C2 protein sequences were obtained. A first set of sequences (#72/41) had a N-terminal  $C_{H\mu}$  region identical to the one obtained by protein sequencing of the 28 Kd papain fragments of the  $\gamma$ 3 heavy chain, a short hinge region containing 3 cysteines and a variable region corresponding to the framework (FR4) residues encoded by the J minigenes adjoining the hinge. The  $C_{H\mu}$  domain is entirely lacking. This cDNA corresponds to the  $\gamma$ 3 chain (Table 4).

In one closely related sequence (#72/1) the proline in position 259 is replaced by threonine.

The sequence corresponding to the  $C_{H\mu}$ 3 and the remaining part of the  $C_{H\mu}$ 2 was obtained by PCR of the cDNA using as KpnI primer a poly T in which a KpnI restriction site had been inserted at the 5' end. The total sequence of the  $\gamma$ 3 chain corresponds to a molecular weight (MW) which is in good agreement with the data obtained from SDS PAGE electrophoresis.

The sequence of this  $\gamma$ 3 chain presents similarities with other  $\gamma$  chains except that it lacks the  $C_{H\mu}$  domain, the  $V_{H\mu}$  domain being adjacent to the hinge.

One or all three of the cysteines could be probably responsible for holding the two  $\gamma$ 3 chains together.

These results have allowed us to define a model for the IgG3 molecule based on sequence and papain cleavage (FIG. 5).

Papain can cleave the molecule on each side of the hinge disulfides and also between  $C_{H\mu}$ 2 and  $C_{H\mu}$ 3. Under non reducing conditions the  $V_{H\mu}$  domains of IgG3 can be isolated as disulfide linked dimer or as monomer depending on the site of papain cleavage.

A second set of clones #72/29 had a slightly different sequence for the  $C_{H\mu}$ 2 and was characterized by a very long hinge immediately preceded by the variable domain. This hinge region has 3 cysteines at its C-terminal end and in a sequence homologous to the  $\gamma$ 3 hinge. Such second set of clones could represent the IgG2 subclass. For the constant part of the  $\gamma$ 3 and also for the putative  $\gamma$ 2, most clones are identical showing the  $\gamma$ 2 or  $\gamma$ 3 specific sequences. A few clones such as #72/1 however show minor differences. For

instance in the case of clones #72/1 two nucleotide differences are detected.

Several  $V_{H\mu}$  regions cDNA's have now been totally or partially sequenced with the exception of a short stretch at the N-terminal end which is primer derived.

Upon translation the majority shows by the characteristic heavy chain Ser<sub>21</sub> Cys<sub>22</sub> and Tyr<sub>90</sub>, Tyr<sub>91</sub> Cys<sub>92</sub> sequences, of the intra  $V_{H\mu}$  region disulfide bridge linking residues 22 and 92. All these clones have a sequence corresponding to the framework 4 (FR4) residues of the variable region immediately preceding the postulated hinge sequence (Table 3). This sequence is generated by the J minigenes and is in the majority of cases similar to the sequence encoded by the human and murine minigenes. The sequence length between region Cys<sub>92</sub> and the C-terminal end of the  $V_{H\mu}$  regions is variable and, in the sequences determined, range from 25 to 37 amino-acids as one might expect from the rearrangements of J and D minigenes varying in length.

Several important questions are raised by the sole existence of these heavy chain immunoglobulins in a non pathological situation. First of all, are they bona fide antibodies? The heavy chain immunoglobulins obtained from trypanosome infected camels react with a large number of parasite antigens as shown in part I of these examples. This implies that the camelid immune system generates an extensive number of binding sites composed of single  $V_{H\mu}$  domains. This is confirmed by the diversity of the  $V_{H\mu}$  regions of the heavy chain immunoglobulins obtained by PCR.

The second question is "how are they secreted?". The secretion of immunoglobulin heavy chains composing four-chain model immunoglobulins does not occur under normal conditions. A chaperoning protein, the heavy chain binding protein, or BIP protein, prevents heavy chains from being secreted. It is only when the light chain displaces the BIP protein in the endoplasmic reticulum that secretion can occur (13).

The heavy chain dimer found in the serum of human or mice with the so-called "heavy chain disease" lack the  $C_{H\mu}$ 1 domains thought to harbour the BIP site (14). In the absence of thi domain the BIP protein can no longer bind and prevent the transport of the heavy chains.

The presence in camels of a IgG1 class composed of heavy and light chains making up between 25% and 50% of the total IgG molecules also raises the problem as to how maturation and class switching occurs and what the role of the light chain is. The camelid light chain appears unusually large and heterogeneous when examined in SDS PAGE.

The largest dimension of an isolated domain is 40 Å and the maximum attainable span between binding sites of a conventional IgG with  $C_{H\mu}$ 1 and  $V_{H\mu}$  will be of the order of 160 Å ( $2(V_{H\mu}+2C_{H\mu}1)$ ) (19). The deletion of  $C_{H\mu}$ 1 domain in the two types of heavy chain antibodies devoid of light chains, already sequenced has, as a result, a modification of this maximum span (FIG. 6). In the IgG3 the extreme distance between the extremities of the  $V_{H\mu}$  regions will be of the order of 80 Å ( $2(V_{H\mu})$ ). This could be a severe limitation for agglutinating or cross linking. In the IgG2 this is compensated by the extremely long stretch of hinge, composed of a 12-fold repeat of the sequence Pro-X (where X is Gln, Lys or Glu) and located N-terminal to the hinge disulfide bridges. In contrast, in the human IgG3, the very long hinge which also apparently arose as the result of sequence duplication does not contribute to increase the distance spanning the two binding sites as this hinge is inter-spaced with disulfide bridges.

The single  $V_{H\mu}$  domain could also probably allow considerably rotational freedom of the binding site versus the Fc domain.

Unlike myeloma heavy chains which result probably from  $C_{\mu}1$  deletion in a single antibody producing cell, or heavy chain antibodies produced by expression cloning(15), the camelid heavy chain antibodies (devoid of light chains) have emerged in a normal immunological environment and it is expected that they will have undergone the selective refinement in specificity and affinity accompanying B cell maturation.

#### Expression and Purification of the Camel $V_{HH}21$ (DR21 on FIG. 7) Protein from *E. coli*

The clones can be expressed in several types of expression vectors. As an example using a commercially available vector Immuno PBS (Huse et al: Science (1989) 246, 1275), clones produced in Bluescript® according to the above described procedure, have been recovered by PCR using the same XhoI containing 5' primer and a new 3' primer, corresponding to residues 113-103 in the framework of the immunoglobulins, in which an Spe site has been constructed: TC TTA ACT AGT GAG GAG ACG GTG ACC TG (SEQ ID NO:51). This procedure allowed the cloning of the  $V_{HH}$  in the Xho/Spe site of the Immuno PBS vector. However, the 3' end of the gene was not in phase with the identification "tag" and the stop codon of the vector. To achieve this, the construct was cut with Spe and the 4 base overhangs were filled in, using the Klenow fragment after which the vector was religated.

The expression vector plasmid ipBS (immunopBS) (Stratagene) contains a pel B leader sequence which is used for immunoglobulin chain expression in *E. coli* under the promoter pLAC control, a ribosome binding site, and stop codons. In addition, it contains a sequence for a c-terminal decapeptide tag.

*E. coli* JM101 harboring the ipBS- $V_{HH}21$  plasmid was grown in 1 l of TB medium with 100 µg/ml ampicillin

and 0.1 % glucose at 32° C. Expression was induced by the addition of 1 mM IPTG (final concentration) at an  $OD_{550}$  of 1.0. After overnight induction at 28° C., the cells were harvested by centrifugation at 4,000 g for 10 min (4° C.) and resuspended in 10 ml TES buffer (0.2M Tris-HCl pH 8.0, 0.5 mM EDTA, 0.5M sucrose). The suspension was kept on ice for 2 hours. Periplasmic proteins were removed by osmotic shock by addition of 20 ml TES buffer diluted 1:4 v/v with water, kept on ice for one hour and subsequently centrifuged at 12,000 g for 30 min. at 4° C. The supernatant periplasmic fraction was dialysed against Tris-HCl pH 8.8, NaCl 50 mM, applied on a fast Q Sepharose flow (Pharmacia) column, washed with the above buffer prior and eluted with a linear gradient of 50 mM to 1M NaCl in buffer.

Fractions containing the  $V_{HH}$  protein were further purified on a Superdex 75 column (Pharmacia) equilibrated with PBS buffer (0.01M phosphate pH 7.2, 0.15M NaCl). The yield of purified  $V_{HH}$  protein varies from 2 to 5 mg/l cell culture.

Fractions were analyzed by SDS-PAGE(I). Positive identification of the camel  $V_{HH}$  antibody fragment was done by Western Blot analysis using antibody raised in rabbits against purified camel IgGH<sub>3</sub> and an anti-rabbit IgG-alkaline phosphatase conjugate (II).

As protein standards (Pharmacia) periplasmic proteins prepared from 1 ml of IPTG-induced JM101/ipBS  $V_{HH}21$  were used. FIG. 8 shows: C,D:fractions from fast S Sepharose column chromatography (C:Eluted at 650 mM NaCl D:Eluted at 700 mM NaCl) E,F:fractions from Superdex 75 column chromatography.

As can be seen, the major impurity is eliminated by ionexchange chromatography and the bulk of the remaining impurities are eliminated by gel filtration.





TABLE 2

Printer		10	20	30	
Derived		G G S V O T G G S L R L S C E I S G L T F	D #72/4	SEQ ID NO:1	
		G G S V O T G G S L R L S C A V S G F T F S	#72/3	SEQ ID NO:2	
		G G S E O G G G S L R L S C A I S G Y T Y G	#72/7	SEQ ID NO:3	
		G G S V O P G G S L T L S C T V S G A T Y S	#72/17	SEQ ID NO:4	
		G G S V O A G G S L R L S C T G S G F F Y S	#72/18	SEQ ID NO:5	
D V Q L V A S G G G S V G A G G S L R L S C T A S G D S F S		#72/2 SEQ ID NO:58			
E V K L V E S G G G L V E P G G S L R L S C A T S G F T F S		Mouse V <sub>H</sub> III <sub>8</sub> SEQ ID NO: 58			
E V Q L L S G G G L V Q P G G S L R L S C A A S G F T F S		Human V <sub>H</sub> III SEQ ID NO: 119			

TABLE 3

Comparison of some Framework 4 residues found in the Camel V<sub>H</sub>III region with the Framework 4 residues corresponding to the consensus region of the Human and Mouse J<sub>H</sub>III genes.

Frame Work 4		J Genes	
Human	W G O G T L V T V S S	SEQ ID NO:9	J1, J4, J5
	W G R G I L V T V S S	SEQ ID NO:130	J2
	W G O G T T V T V S S	SEQ ID NO:120	J6
	W G Q G T M V T V S S	SEQ ID NO:121	J3
	W G Q G T T L T V S S	SEQ ID NO:122	J1
Murine	W G Q G T L V T V S S	SEQ ID NO:9	J2
	W G Q G T S V T V S A	SEQ ID NO:123	J3
	W G A G T T V T V S S	SEQ ID NO:124	J4
		cDNA Clones	
Camel	W G O G T O V T V S S	SEQ ID NO:8	Clones
	W G O G T O V T V S S	SEQ ID NO:8	# 72/19 -# 72/3
	W G O G T L V T V S S	SEQ ID NO:9	1 Clone
	W G R G T Q V T V S S	SEQ ID NO:59	# 72/24
	W G O G T H V T V S S	SEQ ID NO:60	# 72/21
	W G O G I O V T A S S	SEQ ID NO:61	# 72/16

TABLE 4

		SEQ ID NO:125	MOUSE	V <sub>H</sub> III sequence
95	d y y g s s - - - - - y - - - - - f - - - - - d v	W G A G T T V T V S S		
100	a b c d e f g h i j k - - - - -	101 105 110		
1	a l q p g g y c g y g x - - - - - c l	W G Q G T Q V T V S S	SEQ ID NO:13	
2	v a l m d r i s q h - - - - - g c	R G Q G T Q V T V S L	SEQ ID NO:14	
3	v p e h l g p g a i l d l k k y - - - - - k y	W G Q G T Q V T V S S	SEQ ID NO:15	
4	f c y s t a g d g g s g e - - - - - m y	W G Q G T Q V T V S S	SEQ ID NO:16	
7	e i s g s c e l p l i l f - - - - - d y	W G Q G T Q V T V S S	SEQ ID NO:17	
9	d w k y w t e g s q t g g y f - - - - - g q	W G Q G T Q V T V S S	SEQ ID NO:18	
11	r l t e m g a c d a r w a l t a t r t f a y n y	W G Q G T Q V T V S S	SEQ ID NO:19	Random sample
13	g k k d t r t r w a e p r e w - - - - - n n	W G Q G T Q V T V S S	SEQ ID NO:20	
16	g s r f s s p y g a t s r l e e - s d y - - n y	W G Q G T Q V T V S S	SEQ ID NO:21	
17	a d o s l y y a i l x i e y - - - - - k y	W G Q G T Q V T V S S	SEQ ID NO:22	Different camel
18	d a p c y m p t m p a p p i r d a f g w - - d d	F G Q G T Q V T V S S	SEQ ID NO:23	region
19	i s t y w y c t i a p y - - - - - n y	W G Q G T Q V T V S S	SEQ ID NO:24	
20	t i c i e w y g c n l r t i f - - - - - t r	W G Q G T Q V T V S S	SEQ ID NO:25	

TABLE 4-continued

21 nqlaggwyl d pnywls v g a y - - a i W G Q G T Q V T V S S SEQ ID NO:26  
 24 r l e m g a c d a r w a i l a t r i f a y n y W G R G T Q V T V S S SEQ ID NO:27  
 25 d g w t r k e g g i g l p w s v q c e d g y n y W G Q G T Q V T V S S SEQ ID NO:28  
 27 d a y p c h i l l - - - - - d v W G R G T Q V T V S S SEQ ID NO:29  
 29 v e y p i a d m c s - - - - - r y G D P G T Q V T V S S SEQ ID NO:30

## CDR3

Human ← mouse — size range 0-19 aa over 600 entries.  
 camel 8-24 aa 18 entries

EVQLVESGGG				23	SCAASG	SEQ ID NO:80	CDR1	WVROA	PGKGLEWVS	SEQ ID NO:81	CDR2	
GG				SVQGGGSLRL	SCAISG	SEQ ID NO:84	CDR1	WREFG	PGKEREGIA	SEQ ID NO:85	CDR3	
GG				SVQAGGSLRL	SCASSS	SEQ ID NO:88	CDR1	WYRQA	PGKEREFVS	SEQ ID NO:89	CDR2	
RFTIS				70	RDNSKNTLYL	QMNSLRAEDTAVY	YCAR	SEQ ID NO:82	CDR3	WGQGLTVLT	VSS	SEQ ID NO:83
RFTIS				80	QDSTLKTMYL	LMNNLPKPDITGY	YCAA	SEQ ID NO:86	CDR3	WGQGLTVLT	VSS	SEQ ID NO:91
RFTIS				90	QDSAKNTVYL	QMNSLKPDITAMY	YCKI	SEQ ID NO:90	CDR3	WGQGLTVLT	VSS	SEQ ID NO:87
camel V <sub>H</sub>				hinge	GTNEVCCKPKCP	APELLPGG	PSVTVP	SEQ ID NO:51	C <sub>H</sub> 2			
camel				WGQGLTVLT	VSS	EPKIPQPOPKPOPOP	QPOPKPOP	KPEFECTCPKP	APELLGG	PSVGIFFP	SEQ ID NO:87	
human C <sub>H</sub> 1				hinge	ELKTPLEGDTHTICPSCP	APFVAG	PSVFLP	SEQ ID NO:126	C <sub>H</sub> 2			
human gamma 3				KVDKRV	AEFKSCDKTHTICPSCP	APFVAG	PSVFLP	SEQ ID NO:127	C <sub>H</sub> 2			
human gamma 1				KVDKK	AEFKSDTTPPCPSCP	APFVAG	PSVFLP	SEQ ID NO:128	C <sub>H</sub> 2			
human gamma 2				KVKTV	ERKCCVCPSCP	APFVAG	PSVFLP	SEQ ID NO:129	C <sub>H</sub> 2			
human gamma 4				KVDKRV	ESKYGPFCPSCP	APFVAG	PSVFLP	SEQ ID NO:130	C <sub>H</sub> 2			

## REFERENCES

- Ward, E. S., Gussow, D., Griffiths, A. D., Jones, P. T. and Winter G. Nature 341, 544-546 (1989).
- Ungar-Waron H., Eliase E., Gluckman A. and Trainin Z. Isr. J. Vet. Med., 43, 198-203 (1987).
- Bajanya Songa E. and Hamers R., Ann. Soc. Belge Med. trop., 68, 233-240 (1988).
- Edelman G. M., Olins D. E., Gally J. A. and Zinder N. D., Proc. Nat. Acad. Sci., 50, 753 (1963).
- Frank E. and Neelin R. S., Biokhimiya, 28, 193 (1963).
- Reitt I. M., Brostoff J. and Male D. K., Immunology, 40, 1-10 (1985).
- Schiffer M., Girling R. L., Ely K. R. and Edmundson B., Biochemistry, 12, 4620-4631 (1973).
- Fleischman J. B., Pain R. H. and Porter R. R., Arch. Biochem. Biophys., Suppl. 1, 174 (1962).
- Robott O., Onoue K. and Pressman D., PNAS 51, 173-178 (1964).
- Seligmann M., Mihaesco E., Preud'homme J. L., Danon F. and Brouet J. C., Immunological Rev., 48, 145-167 (1979).
- Hendershot L., Bole D., Kohler G. and Kearney J. F., The Journal of Cell Biology, 104, 761-767 (1987).
- Hendershot L. M., The Journal of Cell Biology, 111, 829-837 (1990).
- Hamers-Casterman, C., E. Wittouck, W. Van der Loo and R. Hamers, Journal of Immunogenetics, 6, 373-381 (1979).

14. Applied Biosystems—Ethanol Precipitation of Electra Eluted Electrolysed Sample. Issue n° 27.  
 15. Maniatis, T. E. F. Fritsch and J. Sambrook, Molecular Cloning. A Laboratory Manual (1988).  
 16. Sastry et al., PNAS, 80, 5728, (1989).  
 17. Sanger, F., S. Nicklen and A. R. Coulson, Proc. Natl. Acad. Sci., U.S.A., 74, 5463–5467 (1977).

18. Kabat E. A., Tai Te Wu, M. Reid-Miller, H. M. Perry and K. S. Gottesman, U.S. Dpt of Health and Human Services, Public Health Service, National Institutes of Health (1987).  
 19. Valentine, R. C. and N. M. Geen, J.M.B., 27, 615–617 (1967).

## SEQUENCE LISTING

## ( 1 ) GENERAL INFORMATION:

( 1.1 ) NUMBER OF SEQUENCES: 130

## ( 2 ) INFORMATION FOR SEQ ID NO:1:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 22 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( 1.1 ) MOLECULE TYPE: peptide

## ( 1.x ) FEATURE:

- ( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1..22  
 ( D ) OTHER INFORMATION: /label=FRAMEWORK 1

## ( x.1 ) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Gly Gly Ser Val Gln Thr Gly Gly Ser Leu Arg Leu Ser Cys Gln Ile
1           5           10           15
Ser Gly Leu Thr Phe Asp
20
```

## ( 2 ) INFORMATION FOR SEQ ID NO:2:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 22 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( 1.1 ) MOLECULE TYPE: peptide

## ( 1.x ) FEATURE:

- ( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1..22  
 ( D ) OTHER INFORMATION: /label=FRAMEWORK 1

## ( x.1 ) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Gly Gly Ser Val Gln Thr Gly Gly Ser Leu Arg Leu Ser Cys Ala Val
1           5           10           15
Ser Gly Phe Ser Phe Ser
20
```

## ( 2 ) INFORMATION FOR SEQ ID NO:3:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 22 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( 1.1 ) MOLECULE TYPE: peptide

## ( 1.x ) FEATURE:

- ( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1..22  
 ( D ) OTHER INFORMATION: /label=FRAMEWORK 1



-continued

( 2 ) INFORMATION FOR SEQ ID NO:7:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 21 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..21
      ( D ) OTHER INFORMATION: /label=FRAMEWORK 1

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:7:

  Gly  Gly  Ser  Val  Gln  Ala  Gly  Gly  Ser  Leu  Arg  Leu  Ser  Cys  Val  Ser
   1          5          10          15

  Phe  Ser  Pro  Ser  Ser
           20

```

( 2 ) INFORMATION FOR SEQ ID NO:8:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..11
      ( D ) OTHER INFORMATION: /label=FRAMEWORK 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:8:

  Trp  Gly  Gln  Gly  Thr  Gln  Val  Thr  Val  Ser  Ser
   1          5          10

```

( 2 ) INFORMATION FOR SEQ ID NO:9:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..11
      ( D ) OTHER INFORMATION: /label=FRAMEWORK 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:9:

  Trp  Gly  Gln  Gly  Thr  Leu  Val  Thr  Val  Ser  Ser
   1          5          10

```

( 2 ) INFORMATION FOR SEQ ID NO:10:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain

```

( B ) LOCATION: L..11  
( D ) OTHER INFORMATION: /labd=FRAMEWORK 4

( π i ) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Trp Gly Gln Gly Ala Gln Val Thr Val Ser Ser  
1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:11:

( 1 ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 11 amino acids
- ( B ) TYPE: amine acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( 1 1 ) MOLECULE TYPE: peptide

( 1 x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: L-11
- ( D ) OTHER INFORMATION: /label=FRAMEWORK 4

( 8 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Trp Gly Glu Gly Thr Gln Val Thr Ala Ser Ser  
1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:12:

( 1 ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 11 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( 1 1 ) MOLECULE TYPE: peptide

( 1 x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: L11
- ( D ) OTHER INFORMATION: /label=FRAMEWORK 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Arg Gly Gln Gly Thr Gln Val Thr Val Ser Leu  
1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:13:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 amino acids
- (B) TYPE: amine acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

( 1 1 ) MOLECULE TYPE: peptide

( 1 x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: L\_14
- ( D ) OTHER INFORMATION: /label=CDR3

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ala Leu Gln Pro Gly Gly Tyr Cys Gly Tyr Gly Xaa Cys Leu Trp Gly  
1 5 10 15  
Gln Gly Thr Gln Val Thr Val Ser Ser  
20 25

( 2 ) INFORMATION FOR SEQ ID NO:14:

( 1 ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 23 amino acids  
( B ) TYPE: amino acid

-continued

```

      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..12
      ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Val Ser Leu Met Asp Arg Ile Ser Gln His Gly Cys Arg Gly Gln Gly
1          5          10          15

Thr Gln Val Thr Val Ser Leu
20

( 2 ) INFORMATION FOR SEQ ID NO:15:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 29 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..18
      ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Val Pro Ala His Leu Gly Pro Gly Ala Ile Leu Asp Leu Lys Lys Tyr
1          5          10          15

Lys Tyr Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser
20          25

( 2 ) INFORMATION FOR SEQ ID NO:16:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 26 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..15
      ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Phe Cys Tyr Ser Thr Ala Gly Asp Gly Gly Ser Gly Glu Met Tyr Trp
1          5          10          15

Gly Gln Gly Thr Gln Val Thr Val Ser Ser
20          25

( 2 ) INFORMATION FOR SEQ ID NO:17:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 26 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..15

```

-continued

( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```

Glu Leu Ser Gly Gly Ser Cys Glu Leu Pro Leu Leu Phe Asp Tyr Trp
1          5          10
Gly Gln Gly Thr Gln Val Thr Val Ser Ser
          20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:18:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 28 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

```

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

```

      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..17
      ( D ) OTHER INFORMATION: /label=CDR3

```

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

Asp Trp Lys Tyr Trp Thr Cys Gly Ala Gln Thr Gly Gly Tyr Phe Gly
1          5          10
Gln Trp Gly Gln Gly Ala Gln Val Thr Val Ser Ser
          20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:19:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 35 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

```

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

```

      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..24
      ( D ) OTHER INFORMATION: /label=CDR3

```

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

Arg Leu Thr Glu Met Gly Ala Cys Asp Ala Arg Trp Ala Thr Leu Ala
1          5          10
Thr Arg Thr Phe Ala Tyr Asn Tyr Trp Gly Gln Gly Thr Gln Val Thr
          20          25
Val Ser Ser
          30          35

```

( 2 ) INFORMATION FOR SEQ ID NO:20:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 27 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

```

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

```

      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..16
      ( D ) OTHER INFORMATION: /label=CDR3

```

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Gln Lys Lys Asp Arg Thr Arg Trp Ala Glu Pro Arg Glu Trp Asn Asn

```



-continued

1	5	10	15
Trp Gly Gln Gly Thr Gln Val Thr Ala Ser Ser			
	20	25	

## ( 7 ) INFORMATION FOR SEQ ID NO:21:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 32 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:  
 ( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1..21  
 ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Gly Ser Arg Phe Ser Ser Pro Val Gly Ser Thr Ser Arg Leu Glu Ser	
1	5 10 15
Ser Asp Tyr Asn Tyr Trp Gly Gln Gly Thr Glx Val Thr Ala Ser Ser	
	20 25 30

## ( 2 ) INFORMATION FOR SEQ ID NO:22:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 27 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:  
 ( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1..16  
 ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Ala Asp Pro Ser Ile Tyr Tyr Ser Ile Leu Xaa Ile Glu Tyr Lys Tyr	
1	5 10 15
Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser	
	20 25

## ( 2 ) INFORMATION FOR SEQ ID NO:23:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 33 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:  
 ( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1..22  
 ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Asp Ser Pro Cys Tyr Met Pro Thr Met Pro Ala Pro Pro Ile Arg Asp	
1	5 10 15
Ser Phe Gly Tyr Asp Asp Phe Gly Gln Gly Thr Gln Val Thr Val Ser	
	20 25 30

Ser

-continued

( 2 ) INFORMATION FOR SEQ ID NO:24:

```

( i ) SEQUENCE CHARACTERISTICS:
  ( A ) LENGTH: 26 amino acids
  ( B ) TYPE: amine acid
  ( C ) STRANDEDNESS: single
  ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
  ( A ) NAME/KEY: Domain
  ( B ) LOCATION: 1..15
  ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:24:

  Thr  Ser  Ser  Phe  Tyr  Trp  Tyr  Cys  Thr  Thr  Ala  Pro  Tyr  Asn  Val  Trp
   1          5          10          15

  Gly  Gln  Gly  Thr  Gln  Val  Thr  Val  Ser  Ser
   20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:25:

```

( i ) SEQUENCE CHARACTERISTICS:
  ( A ) LENGTH: 27 amino acids
  ( B ) TYPE: amine acid
  ( C ) STRANDEDNESS: single
  ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
  ( A ) NAME/KEY: Domain
  ( B ) LOCATION: 1..16
  ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:25:

  Thr  Gln  Ile  Gln  Trp  Tyr  Gly  Cys  Asn  Leu  Arg  Thr  Thr  Phe  Thr  Arg
   1          5          10          15

  Trp  Gly  Gln  Gly  Thr  Gln  Val  Thr  Val  Ser  Ser
   20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:26:

```

( i ) SEQUENCE CHARACTERISTICS:
  ( A ) LENGTH: 33 amino acids
  ( B ) TYPE: amine acid
  ( C ) STRANDEDNESS: single
  ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
  ( A ) NAME/KEY: Domain
  ( B ) LOCATION: 1..22
  ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:26:

  Asn  Gln  Leu  Ala  Gly  Gly  Trp  Tyr  Leu  Asp  Pro  Asn  Tyr  Trp  Leu  Ser
   1          5          10          15

  Val  Gly  Ala  Tyr  Ala  Ile  Trp  Gly  Gln  Gly  Thr  His  Val  Thr  Val  Ser
   20          25          30

  Ser

```

( 2 ) INFORMATION FOR SEQ ID NO:27:

```

( i ) SEQUENCE CHARACTERISTICS:
  ( A ) LENGTH: 35 amino acids
  ( B ) TYPE: amine acid
  ( C ) STRANDEDNESS: single

```

-continued

( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: 1..74
- ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```

A r g   L e u   T h r   G l u   M e t   G l y   A l a   C y s   A s p   A l a   A r g   T r p   A l a   T h r   L e u   A l a
1             5             10             15

T h r   A r g   T h r   P h e           A l a   T y r   A s n   T y r   T r p           G l y   A r g   G l y   T h r   G l n   V a l   T h r
                20           25           30

V a l   S e r   S e r
35

```

( 2 ) INFORMATION FOR SEQ ID NO:28:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 35 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: 1..24
- ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```

A s p   G l y   T r p   T h r   A r g   L y s   G l u   G l y   G l y   I l e   G l y   L e u   P r o   T r p   S e r   V a l
1             5             10             15

G l n   C y s   G l u   A s p           G l y   T y r   A s n   T y r   T r p           G l y   G l n   G l y   T h r   G l n   V a l   T h r
                20           25           30

V a l   S e r   S e r
35

```

( 2 ) INFORMATION FOR SEQ ID NO:29:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 21 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: 1..10
- ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```

A s p   S e r   T y r   P r o           C y s   H i s   L e u   L e u   A s p   V a l   T r p   G l y   G l n   G l y   T h r   G l n
1             5             10             15

V a l   T h r   V a l   S e r   S e r
                20

```

( 2 ) INFORMATION FOR SEQ ID NO:30:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 23 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

-continued

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: 1..12
- ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```

Val  Glu  Tyr  Pro  Ile  Ala  Asp  Met  Cys  Ser  Arg  Tyr  Gly  Asp  Pro  Gly
 1          5          10          15

Thr  Gln  Val  Thr  Val  Ser  Ser
      20

```

( 2 ) INFORMATION FOR SEQ ID NO:31:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 27 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: 1..27
- ( D ) OTHER INFORMATION: /label=CH2

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```

Ala  Pro  Glu  Leu  Leu  Gly  Gly  Pro  Thr  Val  Phe  Ile  Phe  Pro  Pro  Lys
 1          5          10          15

Pro  Lys  Asp  Val  Leu  Ser  Ile  Thr  Leu  Thr  Pro
      20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:32:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 27 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: 1..27
- ( D ) OTHER INFORMATION: /label=CH2

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```

Ala  Pro  Glu  Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Thr  Lys
 1          5          10          15

Pro  Lys  Asp  Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro
      20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:33:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 27 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

- ( A ) NAME/KEY: Domain
- ( B ) LOCATION: 1..27
- ( D ) OTHER INFORMATION: /label=CH2

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:33:

-continued

```

Ala Pro Glu Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro Pro Lys
1          5          10          15

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro
20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:34:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 27 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..27
      ( D ) OTHER INFORMATION: /label=CH2

```

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro Pro Lys
1          5          10          15

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro
20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:35:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 12 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..12
      ( D ) OTHER INFORMATION: /label=CH3

```

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```

Gly Glu Thr Arg Glu Pro Glu Val Tyr Thr Leu Ala
1          5          10

```

( 2 ) INFORMATION FOR SEQ ID NO:36:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 18 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..18
      ( D ) OTHER INFORMATION: /label=CH3

```

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```

Gly Glu Thr Arg Glu Pro Glu Val Tyr Thr Leu Ala Pro Asa Arg Leu
1          5          10          15

Glu Leu

```

( 2 ) INFORMATION FOR SEQ ID NO:37:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 12 amino acids

```

-continued

( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

( A ) NAME/KEY: Region  
 ( B ) LOCATION: L:12  
 ( D ) OTHER INFORMATION: /label=hinge

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Gly Thr Asn Glu Val Cys Lys Cys Pro Lys Cys Pro  
 1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:38:

( i ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 35 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

( A ) NAME/KEY: Region  
 ( B ) LOCATION: L:35  
 ( D ) OTHER INFORMATION: /label=hinge

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Glu Pro Lys Ile Pro Glu Pro Glu Pro Lys Pro Glu Pro Glu Pro Glu  
 1 5 10 15  
 Pro Glu Pro Lys Pro Glu Pro Lys Pro Glu Pro Glu Cys Thr Cys Pro  
 20 25 30  
 Lys Cys Pro  
 35

( 2 ) INFORMATION FOR SEQ ID NO:39:

( i ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 35 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

( A ) NAME/KEY: Domain  
 ( B ) LOCATION: L:28  
 ( D ) OTHER INFORMATION: /label=CH2

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Val Phe Pro Pro Lys  
 1 5 10 15  
 Pro Lys Asp Val Leu Ser Ile Ser Gly Xaa Pro Lys  
 20 25

( 2 ) INFORMATION FOR SEQ ID NO:40:

( i ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 28 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

( A ) NAME/KEY: Domain

-continued

( B ) LOCATION: 1:28  
 ( D ) OTHER INFORMATION: /label=CH2

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```

Ala  Pro  Glu  Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Thr  Lys
 1          5          10          15

Pro  Lys  Asp  Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro  Lys
          20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:41:

( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 28 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1:28  
 ( D ) OTHER INFORMATION: /label=CH2

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```

Ala  Pro  Glu  Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Pro  Lys
 1          5          10          15

Pro  Lys  Asp  Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro  Lys
          20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:42:

( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 28 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:

( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1:28  
 ( D ) OTHER INFORMATION: /label=CH2

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```

Ala  Pro  Glu  Leu  Leu  Gly  Gly  Pro  Ser  Val  Phe  Ile  Phe  Pro  Pro  Lys
 1          5          10          15

Pro  Lys  Asp  Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro  Lys
          20          25

```

( 2 ) INFORMATION FOR SEQ ID NO:43:

( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 31 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```

Val  Thr  Val  Ser  Ser  Gly  Thr  Asn  Glu  Val  Cys  Lys  Cys  Pro  Lys  Cys
 1          5          10          15

Pro  Ala  Pro  Glu  Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro
          20          25          30

```

( 2 ) INFORMATION FOR SEQ ID NO:44:

-continued

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 54 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( 11 ) MOLECULE TYPE: peptide

( ix ) FEATURE:  
 ( A ) NAME/KEY: Region  
 ( B ) LOCATION: 1-54  
 ( D ) OTHER INFORMATION: /label=ligase

( x1 ) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```
Val Thr Val Ser Ser Glu Pro Lys Ile Pro Gln Pro Gln Pro Lys Pro
1          5          10          15
Gln Pro Gln Pro Glu Pro Gln Pro Lys Pro Gln Pro Lys Pro Glu Pro
20          25          30
Glu Cys Thr Cys Pro Lys Cys Pro Ala Pro Gln Leu Leu Gly Gly Pro
35          40          45
Ser Val Phe Ile Phe Pro
50
```

( 2 ) INFORMATION FOR SEQ ID NO:45:

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 14 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( 11 ) MOLECULE TYPE: peptide

( ix ) FEATURE:  
 ( A ) NAME/KEY: Region  
 ( B ) LOCATION: 1-14  
 ( D ) OTHER INFORMATION: /label=ligase

( ix ) FEATURE:  
 ( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1-14  
 ( D ) OTHER INFORMATION: /label=CH2

( x1 ) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```
Ala Pro Gln Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro
1          5          10
```

( 2 ) INFORMATION FOR SEQ ID NO:46:

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 14 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( 11 ) MOLECULE TYPE: peptide

( ix ) FEATURE:  
 ( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1-14  
 ( D ) OTHER INFORMATION: /label=CH2

( x1 ) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
Ala Pro Gln Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro
1          5          10
```

( 2 ) INFORMATION FOR SEQ ID NO:47:

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 21 base pairs



-continued

```

      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i l ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( x l ) SEQUENCE DESCRIPTION: SEQ ID NO:47:
CGCCATCAAG  GTAACAGTTG  A                                2 1

( 2 ) INFORMATION FOR SEQ ID NO:48:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 22 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i l ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( i x ) FEATURE:
      ( A ) NAME/KEY: misc_feature
      ( B ) LOCATION: 12..17
      ( D ) OTHER INFORMATION: /label=XbaI site
                               / note= "RESTRICTION SITE"

( x l ) SEQUENCE DESCRIPTION: SEQ ID NO:48:
AGGTCGAGCT  GCTCGAGTCT  GG                                2 2

( 2 ) INFORMATION FOR SEQ ID NO:49:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 22 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i l ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( i x ) FEATURE:
      ( A ) NAME/KEY: misc_feature
      ( B ) LOCATION: 12..17
      ( D ) OTHER INFORMATION: /label=XbaI site
                               / note= "Restriction site"

( x l ) SEQUENCE DESCRIPTION: SEQ ID NO:49:
AGCTCCAGCT  GCTCGAGTCT  GG                                2 2

( 2 ) INFORMATION FOR SEQ ID NO:50:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 22 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i l ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( i x ) FEATURE:
      ( A ) NAME/KEY: misc_feature
      ( B ) LOCATION: 12..17
      ( D ) OTHER INFORMATION: /label=XbaI site
                               / note= "restriction site"

( x l ) SEQUENCE DESCRIPTION: SEQ ID NO:50:
AGGTCGAGCT  TCTCGAGTCT  GG                                2 2

( 2 ) INFORMATION FOR SEQ ID NO:51:

```

-continued

---

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 28 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:51:
TCTTAAGTAG  TGAAGAGACG  GTGACCTG                               28

( 2 ) INFORMATION FOR SEQ ID NO:52:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 30 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( i x ) FEATURE:
      ( A ) NAME/KEY: misc_feature
      ( B ) LOCATION: 1..5
      ( D ) OTHER INFORMATION: /label=5pe1

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:52:
CTAGTGACCA  ACCATCACCA  TCACTAATAG                               30

( 2 ) INFORMATION FOR SEQ ID NO:53:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 30 base pairs
      ( B ) TYPE: nucleic acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: Other
      ( A ) DESCRIPTION: DNA (synthetic)

( i x ) FEATURE:
      ( A ) NAME/KEY: misc_feature
      ( B ) LOCATION: 1..30
      ( D ) OTHER INFORMATION: /cotton "Sequence complementary to
      SEQ ID NO: 52"

( i x ) FEATURE:
      ( A ) NAME/KEY: misc_feature
      ( B ) LOCATION: 26..30
      ( D ) OTHER INFORMATION: /label=EcoRI

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:53:
ACGTGGTGGT  AGTGGTAGTG  ATTATCTTAA                               30

( 2 ) INFORMATION FOR SEQ ID NO:54:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 43 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( v ) FRAGMENT TYPE: N-terminal

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Camelus dromedarius

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..25

```

-continued

( D ) OTHER INFORMATION: /label=CH2

( i x ) FEATURE:

( A ) NAME/KEY: Domain

( B ) LOCATION: 26..43

( D ) OTHER INFORMATION: /label=CH3

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:54:

```

Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Pro  Lys  Pro  Lys  Asp
1          5          10          15
Val  Leu  Ser  Ile  Xaa  Gly  Xaa  Pro  Lys  Gly  Glu  Thr  Arg  Glu  Pro  Glu
20          25          30
Val  Tyr  Thr  Leu  Ala  Pro  Xaa  Arg  Leu  Glu  Leu
35          40

```

( 2 ) INFORMATION FOR SEQ ID NO:55:

( i ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 24 amino acids

( B ) TYPE: amide acid

( C ) STRANDEDNESS: single

( D ) TOPOLOGY: linear

( i 1 ) MOLECULE TYPE: peptide

( v ) FRAGMENT TYPE: N-terminal

( v 1 ) ORIGINAL SOURCE:

( A ) ORGANISM: *Canis lupus familiaris*

( i x ) FEATURE:

( A ) NAME/KEY: Domain

( B ) LOCATION: 1..24

( D ) OTHER INFORMATION: /label=CH2  
/ note= "Clone #721"

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```

Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Thr  Lys  Pro  Lys  Asp
1          5          10          15
Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro
20

```

( 2 ) INFORMATION FOR SEQ ID NO:56:

( i ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 24 amino acids

( B ) TYPE: amide acid

( C ) STRANDEDNESS: single

( D ) TOPOLOGY: linear

( i 1 ) MOLECULE TYPE: peptide

( v ) FRAGMENT TYPE: N-terminal

( v 1 ) ORIGINAL SOURCE:

( A ) ORGANISM: *Canis lupus familiaris*

( i x ) FEATURE:

( A ) NAME/KEY: Domain

( B ) LOCATION: 1..24

( D ) OTHER INFORMATION: /label=CH2

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```

Leu  Pro  Gly  Gly  Pro  Ser  Val  Phe  Val  Phe  Pro  Pro  Lys  Pro  Lys  Asp
1          5          10          15
Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro
20

```

( 2 ) INFORMATION FOR SEQ ID NO:57:

-continued

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 24 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( v ) FRAGMENT TYPE: N-terminal

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Camelus dromedarius

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..24
      ( D ) OTHER INFORMATION: /label=CH2

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Leu  Leu  Gly  Gly  Pro  Ser  Val  Phe  Ile  Phe  Pro  Pro  Lys  Pro  Lys  Asp
1          5          10          15

Val  Leu  Ser  Ile  Ser  Gly  Arg  Pro
20

( 2 ) INFORMATION FOR SEQ ID NO:58:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 30 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( v ) FRAGMENT TYPE: N-terminal

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Camelus dromedarius

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..30
      ( D ) OTHER INFORMATION: /label=Framework 1
           / note= "CAMEL"

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Asp  Val  Gln  Leu  Val  Ala  Ser  Gly  Gly  Gly  Ser  Val  Gly  Ala  Gly  Gly
1          5          10          15

Ser  Leu  Arg  Leu  Ser  Cys  Thr  Ala  Ser  Gly  Asp  Ser  Phe  Ser
20          25          30

( 2 ) INFORMATION FOR SEQ ID NO:59:

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Camelus dromedarius

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..11
      ( D ) OTHER INFORMATION: /label=Framework 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Trp  Gly  Arg  Gly  Thr  Gln  Val  Thr  Val  Ser  Ser
1          5          10

```

-continued

## ( 2 ) INFORMATION FOR SEQ ID NO60:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( v i ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Carnius domochinus

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..11
      ( D ) OTHER INFORMATION: /label=Framework 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO60:

  T r p   G l y   G l n   G l y   T h r   H i s   V a l   T h r   V a l   S e r   S e r
   1                               5                               10

```

## ( 2 ) INFORMATION FOR SEQ ID NO61:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 11 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..11
      ( D ) OTHER INFORMATION: /label=Framework 4

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO61:

  T r p   G l y   G l n   G l y   I l e   G l n   V a l   T h r   A l a   S e r   S e r
   1                               5                               10

```

## ( 2 ) INFORMATION FOR SEQ ID NO62:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 14 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1..14
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..14
      ( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO62:

  A l a   L e u   G l n   P r o   G l y   G l y   T y r   C y s   G l y   T y r   G l y   X a a   C y s   L e u
   1                               5                               10

```

## ( 2 ) INFORMATION FOR SEQ ID NO63:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 12 amino acids
      ( B ) TYPE: amino acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

```

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

( 1 x ) FEATURE:

- ( A ) NAME/KEY: Region
- ( B ) LOCATION: L..18
- ( D ) OTHER INFORMATION: /label=VH

Val Pro Ala His Leu Gly Pro Gly Ala Ile Leu Asp Leu Lys Lys Tyr  
1 5 10 15

Lys Tyr

( 1 ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 15 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

```
( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: L15
      ( D ) OTHER INFORMATION: /label=VH
```

Phe Cys Tyr Ser Thr Ala Gly Asp Gly Gly Ser Gly Glu Met Tyr  
1 5 10 15

( 2 ) INFORMATION FOR SEQ ID NO:66:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 15 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i l ) MOLECULE TYPE: protein

( v l ) ORIGINAL SOURCE:
      ( A ) ORGANISM: Carnius domuncibus

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1..15
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..15
      ( D ) OTHER INFORMATION: /label=CDR3

```

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:66:

```

Glu Leu Ser Gly Gly Ser Cys Glu Leu Pro Leu Leu Phe Asp Tyr
1           5           10           15

```

( 2 ) INFORMATION FOR SEQ ID NO:67:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 17 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i l ) MOLECULE TYPE: protein

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1..17
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..17
      ( D ) OTHER INFORMATION: /label=CDR3

```

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:67:

```

Asp Trp Lys Tyr Trp Thr Cys Gly Ala Gln Thr Gly Gly Tyr Phe Gly
1           5           10           15

```

G l n

( 2 ) INFORMATION FOR SEQ ID NO:68:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 24 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i l ) MOLECULE TYPE: protein

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1..24
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1..24
      ( D ) OTHER INFORMATION: /label=CDR3

```

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:68:

-continued

```

Arg Leu Thr Glu Met Gly Ala Cys Asp 10 Arg Trp Ala Thr Leu Ala
1
Thr Arg Thr Phe Ala Tyr Asn Tyr
20

```

( 2 ) INFORMATION FOR SEQ ID NO:69:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 16 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1-16
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1-16
      ( D ) OTHER INFORMATION: /label=CDR3

```

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:69:

```

Glu Lys Lys Asp Arg Thr Arg Trp Ala Glu Pro Arg Glu Trp Asn Asn
1
5
10
15

```

( 2 ) INFORMATION FOR SEQ ID NO:70:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 21 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1-21
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1-21
      ( D ) OTHER INFORMATION: /label=CDR3

```

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:70:

```

Gly Ser Arg Phe Ser Ser Pro Val Gly Ser Thr Ser Arg Leu Glu Ser
1
5
10
15
Ser Asp Tyr Asn Tyr
20

```

( 2 ) INFORMATION FOR SEQ ID NO:71:

```

( i ) SEQUENCE CHARACTERISTICS:
      ( A ) LENGTH: 16 amino acids
      ( B ) TYPE: amine acid
      ( C ) STRANDEDNESS: single
      ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:
      ( A ) NAME/KEY: Region
      ( B ) LOCATION: 1-16
      ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:
      ( A ) NAME/KEY: Domain
      ( B ) LOCATION: 1-16

```





-continued

( B ) LOCATION: L:16  
( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:74:

```

Thr  Glu  Ile  Glu  Trp  Tyr  Gly  Cys  Asn  Leu  Arg  Thr  Thr  Phe  Thr  Arg
1          5          10          15

```

( 2 ) INFORMATION FOR SEQ ID NO:75:

( 1 ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 22 amino acids  
( B ) TYPE: amine acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:  
( A ) NAME/KEY: Region  
( B ) LOCATION: L:22  
( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:  
( A ) NAME/KEY: Domain  
( B ) LOCATION: L:22  
( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:75:

```

Asn  Glu  Leu  Ala  Gly  Gly  Trp  Tyr  Leu  Asp  Pro  Asn  Tyr  Trp  Leu  Ser
1          5          10          15
Val  Gly  Ala  Tyr  Ala  Ile
20

```

( 2 ) INFORMATION FOR SEQ ID NO:76:

( 1 ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 24 amino acids  
( B ) TYPE: amine acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:  
( A ) NAME/KEY: Region  
( B ) LOCATION: L:24  
( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:  
( A ) NAME/KEY: Domain  
( B ) LOCATION: L:24  
( D ) OTHER INFORMATION: /label=CDR3

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:76:

```

Arg  Leu  Thr  Glu  Met  Gly  Ala  Cys  Asp  Ala  Arg  Trp  Ala  Thr  Leu  Ala
1          5          10          15
Thr  Arg  Thr  Phe  Ala  Tyr  Asn  Tyr
20

```

( 2 ) INFORMATION FOR SEQ ID NO:77:

( 1 ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 24 amino acids  
( B ) TYPE: amine acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:  
( A ) NAME/KEY: Region  
( B ) LOCATION: L:24

-continued

( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:

( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1..24  
 ( D ) OTHER INFORMATION: /label=CDR3

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Asp Gly Trp Thr Arg Lys Glu Gly Gly Ile Gly Leu Pro Trp Ser Val  
 1 5 10  
 Glu Cys Glu Asp Gly Tyr Asn Tyr  
 20

( 2 ) INFORMATION FOR SEQ ID NO:78:

( i ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 10 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:

( A ) NAME/KEY: Region  
 ( B ) LOCATION: 1..10  
 ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:

( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1..10  
 ( D ) OTHER INFORMATION: /label=CDR3

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Asp Ser Tyr Pro Cys His Leu Leu Asp Val  
 1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:79:

( i ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 12 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i x ) FEATURE:

( A ) NAME/KEY: Region  
 ( B ) LOCATION: 1..12  
 ( D ) OTHER INFORMATION: /label=VH

( i x ) FEATURE:

( A ) NAME/KEY: Domain  
 ( B ) LOCATION: 1..12  
 ( D ) OTHER INFORMATION: /label=CDR3

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Val Glu Tyr Pro Ile Ala Asp Met Cys Ser Arg Tyr  
 1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:80:

( i ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 26 amino acids  
 ( B ) TYPE: amine acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( i v i ) ORIGINAL SOURCE:

( A ) ORGANISM: Camelus dromedarius

-continued

( 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly  
 1 5 10 15  
 Ser Ieu Arg Ieu Ser Cys Ala Ala Ser Gly  
 20 25

( 2 ) INFORMATION FOR SEQ ID NO:81:

( 1 ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 16 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( 11 ) MOLECULE TYPE: protein

( v1 ) ORIGINAL SOURCE:

- ( A ) ORGANISM: *Camelus dromedarius*

( x1 ) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ser  
 1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:82:

( 1 ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 32 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( 11 ) MOLECULE TYPE: protein

( x1 ) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Glu  
 1 5 10 15  
 Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg  
 20 25 30

( 2 ) INFORMATION FOR SEQ ID NO:83:

( 1 ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 37 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( 11 ) MOLECULE TYPE: protein

( x1 ) SEQUENCE DESCRIPTION: SEQ ID NO:83:

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Thr Asn Glu Val  
 1 5 10 15  
 Cys Lys Cys Pro Lys Cys Pro Ala Pro Glu Leu Pro Gly Gly Pro Ser  
 20 25 30  
 Val Phe Val Phe Pro  
 35

( 2 ) INFORMATION FOR SEQ ID NO:84:

( 1 ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 18 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( 11 ) MOLECULE TYPE: protein

-continued

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Gly Gly Ser Val Gln Gly Gly Gly Ser Leu Arg Leu Ser Cys Ala Ile  
 1 5 10 15

Ser Gly

( 2 ) INFORMATION FOR SEQ ID NO:85:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 14 amino acids
- ( B ) TYPE: amine acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:85:

Trp Phe Arg Gln Gly Pro Gly Lys Gln Arg Gln Gly Ile Ala  
 1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:86:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 32 amino acids
- ( B ) TYPE: amine acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Arg Phe Thr Ile Ser Gln Asp Ser Thr Leu Lys Thr Met Tyr Leu Leu  
 1 5 10 15

Met Asn Asn Leu Lys Pro Gln Asp Thr Gly Thr Tyr Tyr Cys Ala Ala  
 20 25 30

( 2 ) INFORMATION FOR SEQ ID NO:87:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 60 amino acids
- ( B ) TYPE: amine acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:87:

Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser Glu Pro Lys Ile Pro  
 1 5 10 15

Gln Pro Gln Pro Lys Pro Gln Pro Gln Pro Gln Pro Lys Pro  
 20 25 30

Gln Pro Lys Pro Gln Pro Gln Cys Thr Cys Pro Lys Cys Pro Ala Pro  
 35 40 45

Glu Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro  
 50 55 60

( 2 ) INFORMATION FOR SEQ ID NO:88:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 18 amino acids
- ( B ) TYPE: amine acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:88:

-continued

Gly	Gly	Ser	Val	Gln	Ala	Gly	Gly	Ser	Leu	Arg	Leu	Ser	Cys	Ala	Ser
1				5					10					15	
Ser Ser															

( 2 ) INFORMATION FOR SEQ ID NO:89:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 14 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:89:

Trp	Tyr	Arg	Gln	Ala	Pro	Gly	Lys	Glu	Arg	Glu	Phe	Val	Ser
1				5					10				

( 2 ) INFORMATION FOR SEQ ID NO:90:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 37 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:90:

Arg	Phe	Ile	Ser	Gln	Asp	Ser	Ala	Lys	Asn	Thr	Val	Tyr	Leu	Gln	
1				5				10					15		
Met	Asn	Ser	Leu	Lys	Pro	Glu	Asp	Thr	Ala	Met	Tyr	Tyr	Cys	Lys	Ile
			20				25						30		

( 2 ) INFORMATION FOR SEQ ID NO:91:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 37 amino acids  
 ( B ) TYPE: amino acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:91:

Trp	Gly	Gln	Gly	Thr	Gln	Val	Thr	Val	Ser	Ser	Gly	Thr	Asa	Glu	Val
1				5				10						15	
Cys	Lys	Cys	Pro	Lys	Cys	Pro	Ala	Pro	Glu	Leu	Pro	Gly	Gly	Pro	Ser
			20				25						30		
Val	Phe	Val	Phe	Pro											
			35												

( 2 ) INFORMATION FOR SEQ ID NO:92:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 400 base pairs  
 ( B ) TYPE: nucleic acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:92:

CTCGAGTCTG	GGGGAGGATC	GGTGCAGGCT	GGAGGGTCTC	TGAGACTCTC	GTGCGCAGCC	60
TCTGGATACA	GTAATTGTCC	CCTCACTTGG	AGCTGGTATC	GCCAGTTTCC	AGGAACGGAG	120
CGCGAGTTTC	TCTCCAGTAT	GGATCCGGAT	GGAATACCA	AGTACACATA	CTCCGTGAAG	180

-continued

GCCGCTTCA	CCATGTCCCG	AGGCAGCACC	GAGTACACAG	TATTTCTGCA	AATGGACAAT	240
CTGAAACCTG	AGGACACGGC	GATGTATTAC	TGTAAAACAG	CCCTACAACC	TGGGGTTTAT	300
TGTGGGTATG	GGTANTGCCT	CTGGGCCAG	GGGACCCAGG	TCACCGTCTC	CTCACTAGTT	360
ACCCGTACGA	CGTTCCGGAC	TACGGTTCTT	AATAGAATTC			400

( 2 ) INFORMATION FOR SEQ ID NO:93:

( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 391 base pairs

( B ) TYPE: nucleic acid

( C ) STRANDEDNESS: single

( D ) TOPOLOGY: linear

( 1.1 ) MOLECULE TYPE: cDNA

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:93:

CTCGAGTCTG	GGGGAGGGCT	GGTGCAGGCT	GGAGGGTCTC	TGAGACTCTC	CTGTGCATCT	60
TCTTCTAAAT	ATATGCGCTTG	CACCTACGAC	ATGACCTGGT	ACCGCCAGGC	TCCAGGCCAAG	120
GAGCGCGAAT	TTGTCTCAAG	TATAAATATT	GATGCTAAGA	CAACATACGC	AGACTCCGTG	180
AAGGGCCGAT	TCACCATCTC	CCAAGACAGC	GCCAGGAACA	CGGTGTATCT	GCAGATGAAC	240
AGCCTGA AAC	CTGAGGACAC	GGCGATTGAT	TACTGTAAAA	TAGATTCTGA	CCCGTGCCAT	300
CTCCTTGATG	TCTGGGGCCA	GGGGACCCAG	GTCAACCTCT	CCTCACTAGT	TACCCGTACG	360
AGCTTCCGGA	CTACGGTCTT	TAAATAGAAAT	C			391

( 2 ) INFORMATION FOR SEQ ID NO:94:

( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 443 base pairs

( B ) TYPE: nucleic acid

( C ) STRANDEDNESS: single

( D ) TOPOLOGY: linear

( 1.1 ) MOLECULE TYPE: cDNA

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:94:

CAGGTGAAAC	TGCTCGAGTC	TGGAGGAAGC	TGGGTGCAGA	CTGGAGGATC	TCTGAGACTC	60
TCCTGTGCAG	TCTCTGGATT	CTCCTTTAGT	ACCAAGTTGTA	TGGCCTGGTT	CCGCCAGGCT	120
TCAGGAAAGC	AGCGTGAGGG	GTCGCGAGCC	ATTAATAGTG	CGGGTGGTAG	GACATACTAC	180
AACACATATG	TCGCCGAGTC	CGTGAAAGGC	CGATTGCGCA	TCTCCCAAGA	CAACGCCAAG	240
ACCACGGTAT	ATCTTGATAT	GAACAACCTA	ACCCCTGAAG	ACACGGGTAC	GTATTACTGT	300
CGCGCGGTCC	CAGCCCACTT	GGGACCTGGC	GCCATTCTTG	ATTTGAAAAA	GTATAAGTAC	360
TGGGGCCAGG	GGACCCAGGT	CACCGTCTCC	TCACTAGCTA	GTTACCCGTA	CGACGTTCCG	420
GACTACGGTT	CTTAATAGAA	TTC				443

( 2 ) INFORMATION FOR SEQ ID NO:95:

( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 433 base pairs

( B ) TYPE: nucleic acid

( C ) STRANDEDNESS: single

( D ) TOPOLOGY: linear

( 1.1 ) MOLECULE TYPE: cDNA

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:95:

CTCGAGTCTG	GGGGAGGGCT	GGTGCAGGCT	GGAGGGTCTC	TGAGACTCTC	CTGTAATGTC	60
------------	------------	------------	------------	------------	------------	----

-continued

TCTGGCTCTC	CCAGTAGTAC	TTATTGCCTG	GGCTGGTTCC	GCCAGGCTCC	AGGGAGGGAG	120
CGTGAGGGGG	TCACAGCGAT	TAACTACTGAT	GGCAGTATCA	TATACGCAGC	CGACTCCGTG	180
AAGGGCCGAT	TCACCATCTC	CCAAGACACC	GCCAAGGAAA	CGGTACATCT	CCAGATGAAC	240
AACCTGCAAC	CTGAGGATAC	GGCCACCTAT	TACTGCGCGG	CAAGACTGAC	GGAGATGGGG	300
GCTTGATGAT	CGAGATGGGC	GACCTTAGCG	ACAAGGACGT	TTGCGTATAA	CTACTGGGGC	360
CGGGGACACC	AGGTACCCGT	CTCTCACTA	GTTACCCGTA	CGACGTTCCG	GACTACGGTT	420
CCTAATAGAA	TTC					433

( 2 ) INFORMATION FOR SEQ ID NO:96:

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 469 base pairs  
 ( B ) TYPE: nucleic acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( 1.1 ) MOLECULE TYPE: cDNA

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:96:

CAGGTGAAAC	TGCTCGAGTC	TGGGGGAGGG	TGGGTGCAGG	CTGGAGGGTC	TCTGAGACTC	60
TCCTGTAAAT	TCTCTGGCTC	TCCAGTAGT	ACTTATTGCC	TGGGCTGGTT	CGGCCAGGCT	120
CCAGGGAAAG	AGCGTGAGGG	GGTCACAGCG	ATTAACTACT	ATGGCAGTGT	CATATACGCA	180
GCCGACTCCG	TGAAGGGCCG	ATTCAACATC	TCCCAAGACA	CCGCCAAGAA	AACGGTATAT	240
CTCCAGATGA	ACAACCTGCA	ACCTGAGGAT	ACGGCCACCT	ATTACTGCGC	GGCAAGACTG	300
ACGGAGATGG	GGGCTTTGTA	TGCGAGATGG	GCGACCTTAG	CGACAAGGAC	GTTTGCATAT	360
AACTACTGGG	GCCGGGGGAC	CCAGGTCACC	GTCCTCTCAC	TAGCTAGTTA	CCCGTACGAC	420
GTTCCGGACT	ACGGTTCCTA	ATAGAATTC				449

( 2 ) INFORMATION FOR SEQ ID NO:97:

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 424 base pairs  
 ( B ) TYPE: nucleic acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear

( 1.1 ) MOLECULE TYPE: cDNA

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CTCGAAGTCT	GAGGAGGCTC	GGCGCAGGCT	GGAGGATCTC	TGAGACTCTC	CTGTGCAGCC	60
CACGGGATTC	CGCTCAATGG	TTACTACATC	GCCTGGTTCC	GTCAAGGCTCC	TGGGAAGGGG	120
CGTGAGGGGG	TGCAACAAT	TAATGGTGGT	CGCGACGTCA	CATACTACGC	CGACTCCGTG	180
ACGGGCCGAT	TTACCATCTC	CCGAGACAGC	CCCAAGAATA	CGGTGTATCT	GCAGATGAAC	240
AGGCTGAACA	CTGAGGACAC	GGCCATCTAC	TTCTGTGCAG	CAGGCTCGCG	TTTTTCTAGT	300
CCTGTTGGGA	CCACTTCTAG	ACTCGAAAGT	AGCGACTATA	ACTATTGGGG	CCAGGGGATC	360
CAGGTCACCG	TCACTCTACT	AGTTACCCGT	ACGACGTTCC	GGACTACGGT	TCTTAATAGA	420
ATTTC						424

( 2 ) INFORMATION FOR SEQ ID NO:98:

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 415 base pairs  
 ( B ) TYPE: nucleic acid  
 ( C ) STRANDEDNESS: single  
 ( D ) TOPOLOGY: linear



-continued

( 1 ) MOLECULE TYPE: cDNA

( 2 ) SEQUENCE DESCRIPTION: SEQ ID NO:98:

CTCGAGTCTG	GAGGAGGCTC	GGTTCAGGCT	GGAGGGTCCC	TTAGACTCTC	CTGTGCAGCC	6 0
TCTGACTACA	CCATCACTGA	TTATTGCATG	GCCTGGTTCC	GCCAGGCTCC	AGGGGAAGGAG	1 2 0
CGTGAAATTG	TCGCAGCGAT	TCAAGTTGTC	CGTAGTGATA	CTCGGCTCAC	AGACTACGCC	1 8 0
CACCTCCGTGA	AGGGACGATT	CACCATCTCC	CAAGGCAACA	CCAAGAACAC	AGTGAATCTG	2 4 0
CAAAATGAACA	GCCTGACACC	TGAGGACACG	GCCATCTACA	GTGTGCGGGC	AACCACTAGT	3 0 0
TTTACTGGT	ACTGCACCAC	GGCGCCTTAT	AACGTCTGGG	GTCAGGGGAC	CCAGGTCAAC	3 6 0
GTCTCCTCAC	TAGTTACCCG	TACGACGTTT	CGGACTACGG	TTCCTTAATG	AATTC	4 1 5

( 2 ) INFORMATION FOR SEQ ID NO:99:

( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 405 base pairs

( B ) TYPE: nucleic acid

( C ) STRANDEDNESS: single

( D ) TOPOLOGY: linear

( 1 ) MOLECULE TYPE: cDNA

( 2 ) SEQUENCE DESCRIPTION: SEQ ID NO:99:

CTCGAGTCTG	GGGGAGGCTC	GGTGCAGGGT	GGAGGGTCTC	TGAGACTCTC	CTGTGCAATC	6 0
1CTGGATACA	CGTACGGTAG	CTTCTGTATG	GGCTGGTTCC	GGCAGGGTTC	AGGCAAGGAA	1 2 0
CGTGAGGGGA	TCGCAACTAT	TCTTAATGGT	GGTACTAACA	CATACTATGC	CGACTCGGTG	1 8 0
AAGGGCCGAT	TCACCATCTC	CCAAGACAGC	ACGTTGAAGA	CGATGTATCT	GCTAATGAAC	2 4 0
AACCTGAAAC	CTGAAGACAC	GGGCACCTAT	TACTGTGCTG	CAGAACTAAG	TGGTGGTAGT	3 0 0
TGTGAATTGC	CTTTGCTATT	TGACTACTGG	GGCCAGGGCA	CCCAAGGTAC	CGTCTCCTCA	3 6 0
CTAGTTACCC	GTACGACGTT	CCGGACTACG	GTTCTTAATA	GAATTC		4 0 6

( 2 ) INFORMATION FOR SEQ ID NO:100:

( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 427 base pairs

( B ) TYPE: nucleic acid

( C ) STRANDEDNESS: single

( D ) TOPOLOGY: linear

( 1 ) MOLECULE TYPE: cDNA

( 2 ) SEQUENCE DESCRIPTION: SEQ ID NO:100:

CTCGAGTCTG	GGGGAGGCTC	GGTGCAGGGT	GGAGGGTCTC	TGAGACTCTC	CTGTACAGGC	6 0
TCTGGATTCC	CCTATAGTAC	CTTCTGTTCTG	GGGTGGTTCC	GCCAGGCTCC	AGGGGAAGGAG	1 2 0
CGTGAGGGGG	TCGCGGGTAT	TAATAGTGCA	GGAGGTAATA	CTTACTATGC	GCAGCCCGTG	1 8 0
AAGGGCCGAT	TCACCATCTC	CCAAGGGAAAT	GCCAAGAAAT	CGGTGTTTCT	GCAAAATGGAT	2 4 0
AACCTGAAAC	CTGAGGACAC	GGCCATCTAT	TACTGCGCGG	GGGATAGTCC	ATGTTACATG	3 0 0
CCGACTATGC	CCGCTCCCCC	GATACGAGAC	AGTTTGGCT	GGGATGAATT	TGGCCAGGGG	3 6 0
ACCCAGGICA	CCGTCCTTCC	ACTAGTAC	CGTACGACGT	TCCGGACATC	GCTTCTTAAT	4 2 0
AGAATTC						4 2 7

( 2 ) INFORMATION FOR SEQ ID NO:101:

( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 408 base pairs

-continued

( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( 1 ) MOLECULE TYPE: cDNA

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:101:

```
CTCGAGTCAO GGGGAGGCTC GGTACAGGTT GAGGGGTCTC TGAGACTCTC CTGTGTAGCC 60
TCTACTCACA CCGACAGTAG CACCTGTATA GGCTGGTTCC GCCAGGCTCC AGGGAAGGAG 120
CCCGAGGGGG TCGCAAGTAT ATATTTTGGT GATGGTGGTA CGAATTATCG CAGCTCCGTC 180
AAGGGCCGAT TCACCATCTC CCAACTCAAC GCCCAGAAAC CAGTGATCTC GCAATGAAC 240
AGCCTGAAAC CTGAGGACAG CGCCATGTAC TACTGTGCAA TCACTGAAAT TGAGTGGTAT 300
GGGTGCAATT TAAGGACTAC TTTTACTCGC TGGGGCCAGG GGACCCAGGT CACCGTCTCC 360
TCACTAGTTA CCCGTACGAC GTTCCGGACT ACGGTTCTTA ATAGAATTG 409
```

( 2 ) INFORMATION FOR SEQ ID NO:102:

( 1 ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 445 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( 1 ) MOLECULE TYPE: cDNA

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:102:

```
CTCGAGTCTG GGGGAGGCTC GGTACAAACT GGAGGGTCTC TGAGACTCTC TTGCGAAATC 60
TCTGGATTGA CTTTGTATGA TTCTGACGTG GGGTGGTACC GCCAGGCTCC AGGGGATGAG 120
TGCAAAATTGG TCTCAGGTAT TCTGAGTGAT GGTACTCCAT ATACAAAGAG TGGAGACTAT 180
GCTGAGTCTG TGAGGGGCGG GGTACCATC TCCAGAGACA ACGCCAAGAA CATGATATAC 240
CTTCAAAATGA ACGACTGAA ACCTGAGGAC ACGGCCATGT ATTACTGCGC GGTAGATGGT 300
TGGACCCGGA AGGAAGGGGG AATCGGGTTA CCCTGGTCCG TCCAATGTGA AGATGGTTAT 360
AACTATTGGG GCCAGGGGAC CCAGGTCACC GTCTCCTCAC TAGTTACCCG TACGACGTTT 420
CGGACTACGG TTTTAATAG AATTG 445
```

( 2 ) INFORMATION FOR SEQ ID NO:103:

( 1 ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 394 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( 1 ) MOLECULE TYPE: cDNA

( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:103:

```
CTCGAGTCTG GAGGAGGCTC GGTGACAGGT GGAGGGTCTC TGAGACTCTC CTGTGTAGCC 60
TCTGGATTGA ATTTCGAAAC TTCTCGTATG GCGTGGTACC GCCAGACTCC AGGAAATGTG 120
TGTGAGTTGG TCTCAAGTAT TTACAGTGAT GGCAAAACAT ACTACGTGGA CCGCATGAAG 180
GGCCGATTCA CATTCTTAG AGAGAATGCC AAGAATACAT TGATATCACA ACTGAGCGGC 240
CTCAAACTTG AGGACACGGC CATGTATTAC TGTGCGCCGG TTGAATATCC TATTGCAGAC 300
ATGTGTTCCA GATACGGCGA CCGGGGAC CAGGTACCCG TCTCCTCACT AGTTACCCGT 360
ACGACGAACC GGACTACGGT TCTTAATAGA ATTC 394
```

( 2 ) INFORMATION FOR SEQ ID NO:104:

-continued

## ( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 433 base pairs

( B ) TYPE: nucleic acid

( C ) STRANDEDNESS: single

( D ) TOPOLOGY: linear

## ( 1.1 ) MOLECULE TYPE: cDNA

## ( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:101:

```

CTCGAGTCTG GGGGAGGCTC GGTGCAGGTT GGAGGGTCTC TGAAACTCTC CTGTAAAAATC      60
TCTGGAGGTA CCCAGATCG TGTTCCTAAA TCTTTGGCCT GGTTCGCCCA GGCTCCAGAG      120
AAGGAGCGCG AGGGGATCCG AGTTCCTTCG ACTAAGGATG GTAAGACATT CTATGCCGAC      180
TCCGTGAAGG GCCGATTAC CACTCTCTTA GATAATGACA AGACCACCTT CTCCTTACAA      240
CTTGATCGAT TGAACCCGGA GGACACTGCC GACTACTACT GCGCTGCAAA TCAATTAGCT      300
GGTGGCTGGT ATTTGGACCC GAATTACTGG CTCTCTGTGG GTGCATATGC CATCTGGGGC      360
CAGGGGACCC AGGTCACCGT CTCCTCACTA GTTACCCGTA CGACGTTCCG GACTACGGTT      420
CTTAATAGAA TTC                                     433

```

## ( 2 ) INFORMATION FOR SEQ ID NO:105:

## ( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 416 base pairs

( B ) TYPE: nucleic acid

( C ) STRANDEDNESS: single

( D ) TOPOLOGY: linear

## ( 1.1 ) MOLECULE TYPE: cDNA

## ( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:105:

```

CAGGTGAAAC TGCTCGAGTC TGGGGGAGGC TCGGTGCAGG CTGGGGGGGC TCTGACACTC      60
TCTTTGTGAT ACACCAACGA TACTGGGACC ATGGGATGGT TTCGCCAGGC TCCAGGGAAA      120
GAGTGCAGAA GGGTCGCGCA TATTACGCTC GATGGTATGA CCTTCATTGA TGAACCCGCTG      180
AAGGGGCGAT TCACGATCTC CCGAGACAAC GCCCAGAAAA CGTTGTCTTT GCGAATGAAT      240
AGTCTGAGGC CTGAGGACAC GGCCGTGTAT TACTGTGCGG CAGATTGGAA ATACTGGAAT      300
TGTGGTGCCC AGACTGGAAG ATACTTCGGA CAGTGGGGTC AGGGGGCCCA GGTACCCGTC      360
TCCTCACTAG CTAGTTACCC GTACGACGTT CCGGACTACG GTTCTTAATA GAATTC       416

```

## ( 2 ) INFORMATION FOR SEQ ID NO:106:

## ( 1 ) SEQUENCE CHARACTERISTICS:

( A ) LENGTH: 361 base pairs

( B ) TYPE: nucleic acid

( C ) STRANDEDNESS: single

( D ) TOPOLOGY: linear

## ( 1.1 ) MOLECULE TYPE: cDNA

## ( x 1 ) SEQUENCE DESCRIPTION: SEQ ID NO:106:

```

CTCGAGTCTG GGGGAGGCTC GGTCCAACTT GGAGGATCTC TGACACTCTC CTGTACAGTT      60
TCTGGGGCCA CCTACAGTGA CTACAGTATT GGATGGATCC GCCAGGCTCC AGGGAAGGAC      120
CGTGAAGTGA TCGCAGCCGC TAAATACGGT GCGACIAGTA AATTCACGTT CGACTTTTGG      180
AAGGGCCGAT TCACCATTTT CCAAGACAAC GCCAAGAATA CGGTATATCT GCAAATGAGC      240
TTCTTGAAC CTGAGGACAC GGCCATCTAT TACTGTGCGG CAGCGGACCC AAGTATATAT      300
TATAGTATCC TCCATTGAGT ATAAGTACTG GGGCCAGGGG ACCCAGGTCA CGGTCTCTCT      360

```

A

361

-continued

## ( 2 ) INFORMATION FOR SEQ ID NO:107:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 354 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

## ( i i ) MOLECULE TYPE: cDNA

## ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:107:

```

CTC GAG TCA G GGG GAG GCT C GGT GGA GGT C GGA GGG TCT C TGAG ACT CTC CTGTACAGCC      60
TCTGGATACG TATCCTCTAT GGCTTGTTT C GCGCAGGTT C CAGGGCAGGA GCGCGAGGGG      120
CTCGCGTTTG TTCAAACGGC TGACAATAGT GCATTATATG GCGACTCCGT GAAGGGCCGA      180
TTCACCATCT CCCACGACAA CGCCAAGAAC ACCTGTATC TGCAAAATGCG CAACCTGCAG      240
CCTGACGACA CTGGCGTGTA CTACTGTGCG GCCCAAAAGA AGGATCGTAC TAGATGGGCC      300
GAGCCTCGAG AATGGAAACA CTGGGGCCAG GGGACCCAGG TCACCGTCTC CTCA      354

```

## ( 2 ) INFORMATION FOR SEQ ID NO:108:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 381 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

## ( i i ) MOLECULE TYPE: cDNA

## ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:108:

```

CTCGAGTCA G GTGTCGGTC TGATGTGCAG CTGGTGCGT CTGGGGGAG CTCGGTGCAG      60
GCTGGAGGCT CTCTGAGACT CTCCTGTACA GCCTCTGGAG ACAGTTTCAG TAGATTTGCC      120
ATGCTCTGGT TCCGCCAGGC TCCAGGGAAG GAGTGCGAAT TGGTCTCAAG CATTCTAAAGT      180
AATCGAAGGA CAACTGAGGC CGATTCCGTG CAAGGCCGAT TCACCATCTC CCGAGACAAT      240
TCCAGGAACA CAGTGTATCT GCAAAATGAAC AGCCTGAAAC CCGAGGACAC GGCCTGTAT      300
TACTGTGGGG CAGTCTCCCT AATGGACCGA ATTTCCCAAC ATGGGTGCCG GGGCCAGGGA      360
ACCCAGGTCA CCGTCTCCTT A      381

```

## ( 2 ) INFORMATION FOR SEQ ID NO:109:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 18 amino acids
- ( B ) TYPE: amino acid
- ( D ) TOPOLOGY: linear

## ( i i ) MOLECULE TYPE: peptide

## ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:109:

```

Gly Gln Pro Arg Gln Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp
1           5           10           15
Gln Leu

```

## ( 2 ) INFORMATION FOR SEQ ID NO:110:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 18 amino acids
- ( B ) TYPE: amino acid
- ( D ) TOPOLOGY: linear

## ( i i ) MOLECULE TYPE: peptide

-continued

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:110:

Gly	Gln	Pro	Arg	Gln	Pro	Gln	Val	Tyr	Thr	Leu	Pro	Pro	Ser	Arg	Glu
1				5					10					15	
Glu Met															

( 2 ) INFORMATION FOR SEQ ID NO:111:

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 18 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:111:

Gly	Gln	Pro	Arg	Gln	Pro	Gln	Val	Tyr	Thr	Leu	Pro	Pro	Ser	Gln	Glu
1				5					10					15	
Glu Met															

( 2 ) INFORMATION FOR SEQ ID NO:112:

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 24 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp
1				5					10					15	
Thr Leu Met Ile Ser Arg Thr Pro															
				20											

( 2 ) INFORMATION FOR SEQ ID NO:113:

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 23 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:113:

Val	Ala	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp	Thr
1				5					10					15	
Leu Met Ile Ser Arg Thr Pro															
				20											

( 2 ) INFORMATION FOR SEQ ID NO:114:

( 1 ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 24 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:114:

Phe	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp
1				5					10					15	
Thr Leu Met Ile Ser Arg Thr Pro															
				20											

( 2 ) INFORMATION FOR SEQ ID NO:115:

-continued

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 19 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:115:

Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg  
 1 5 10 15  
 Asp Glu Leu

( 2 ) INFORMATION FOR SEQ ID NO:116:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 19 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:116:

Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg  
 1 5 10 15  
 Glu Glu Met

( 2 ) INFORMATION FOR SEQ ID NO:117:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 19 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:117:

Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln  
 1 5 10 15  
 Glu Glu Met

( 2 ) INFORMATION FOR SEQ ID NO:118:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 30 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:118:

Glu Val Lys Leu Val Glu Ser Gly Gly Gly Leu Val Glu Pro Gly Gly  
 1 5 10 15  
 Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Phe Thr Phe Ser  
 20 25 30

( 2 ) INFORMATION FOR SEQ ID NO:119:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 29 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:119:

Glu Val Glu Leu Leu Ser Gly Gly Gly Leu Val Glu Pro Gly Gly Ser

-continued

1	5	10	15
Leu Arg Leu Ser	Cys Ala Ala Ser Gly Phe Thr Phe Ser		
20	25		

( 7 ) INFORMATION FOR SEQ ID NO:120:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 11 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:120:

Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser
1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:121:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 11 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:121:

Trp Gly Gln Gly Thr Met Val Thr Val Ser Ser
1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:122:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 11 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:122:

Trp Gly Gln Gly Thr Thr Leu Thr Val Ser Ser
1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:123:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 11 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:123:

Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ala
1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:124:

( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 11 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:124:

Trp Gly Ala Gly Thr Thr Val Thr Val Ser Ser
1 5 10

## ( 2 ) INFORMATION FOR SEQ ID NO:125:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 21 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:125:

```

A s p   T y r   T y r   G l y   S e r   S e r   T y r   P h e   A s p   V a l   T r p   G l y   A l a   G l y   T h r   T h r
1           5           10           15

V a l   T h r   V a l   S e r   S e r
20

```

## ( 2 ) INFORMATION FOR SEQ ID NO:126:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 67 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:126:

```

L y s   V a l   A s p   L y s   A r g   V a l   G l u   L e u   L y s   T h r   P r o   L e u   G l y   A s p   T h r   T h r
1           5           10           15

H i s   T h r   C y s   P r o   A r g   C y s   P r o   G l u   P r o   L y s   C y s   S e r   A s p   T h r   P r o   P r o
20           25

P r o   C y s   P r o   A r g   C y s   P r o   G l u   P r o   L y s   S e r   C y s   A s p   T h r   P r o   P r o   P r o
35           40           45

C y s   P r o   A r g   C y s   P r o   A l a   P r o   G l u   L e u   L e u   G l y   G l y   P r o   S e r   V a l   P h e
50           55           60

L e u   P h e   P r o
65

```

## ( 2 ) INFORMATION FOR SEQ ID NO:127:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 35 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:127:

```

L y s   V a l   A s p   L y s   L y s   A l a   G l u   P r o   L y s   S e r   C y s   A s p   L y s   T h r   H i s   T h r
1           5           10           15

C y s   P r o   P r o   C y s   P r o   A l a   P r o   G l u   L e u   L e u   G l y   G l y   P r o   S e r   V a l   P h e
20           25           30

L e u   P h e   P r o
35

```

## ( 2 ) INFORMATION FOR SEQ ID NO:128:

- ( i ) SEQUENCE CHARACTERISTICS:  
 ( A ) LENGTH: 31 amino acids  
 ( B ) TYPE: amine acid  
 ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:128:

```

L y s   V a l   L y s   V a l   T h r   V a l   G l u   A r g   L y s   C y s   C y s   V a l   G l u   C y s   P r o   P r o
1           5           10           15

```



-continued

Cys Pro Ala Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro  
20 25 30

(2) INFORMATION FOR SEQ ID NO:129:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 amino acids  
(B) TYPE: amine acid  
(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:129:

Lys Val Asp Lys Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser  
1 5 10 15  
Cys Pro Ala Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro  
20 25 30

(2) INFORMATION FOR SEQ ID NO:130:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 11 amino acids  
(B) TYPE: amine acid  
(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Trp Gly Arg Gly Thr Leu Val Thr Val Ser Ser  
1 5 10

We claim:

1. An immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable ( $V_{Hm}$ ) region and a constant region, said constant region being devoid of first constant domain  $C_{H1}$ , wherein the immunoglobulin is devoid of polypeptide light chains, and wherein the variable region contains in position 45 an amino acid which is not a leucine, proline or glutamine residue.

2. An immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable ( $V_{Hm}$ ) region and a constant region, said constant region being devoid of first constant domain  $C_{H1}$ , wherein the immunoglobulin is devoid of polypeptide light chains, and comprising:

four frameworks in its variable region, which frameworks comprise an amino-acid sequence selected from the following sequences:

for the framework 1 domain

GGSVQTGGSLRLSCFISGLTFD (SEQ ID NO:1)  
GGSVQTGGSLRLSCAVSGFSPS (SEQ ID NO:2)  
GGSEQGGSLRLSCAISGYTYG (SEQ ID NO:3)  
GGSVQPGGSLTSLCTVSGATYS (SEQ ID NO:4)  
GGSVOAGSLRLSCTGSGFPYS (SEQ ID NO:5)  
GGSVOAGSLRLSCVAGFGTS (SEQ ID NO:6)  
GGSVOAGSLRLSCVSFSPSS (SEQ ID NO:7)

for the framework 4 domain

WGQGTQVTYSS (SEQ ID NO:8)  
WGQGLTVTVSS (SEQ ID NO:9)  
WGQGAQVTYSS (SEQ ID NO:10)  
WGQGTQVTYSS (SEQ ID NO:11)  
RQGTQVTYSS (SEQ ID NO:12)

and/or, in that its variable region comprises CDR domains, comprising for the CDR3 domain

ALQPGGYCGYX-----CL (SEQ ID NO:62)  
VSLMDRISQH-----GC (SEQ ID NO:63)  
VPAITLGGAILDLKKY-----KY (SEQ ID NO:64)  
FCYSTAGDGGSGE-----MY (SEQ ID NO:65)  
ELSGGCELPPLF-----DY (SEQ ID NO:66)  
DWKYWTCGAQTGGYF-----GQ (SEQ ID NO:67)  
RLTEMGACDARWATLATRTFAYNY (SEQ ID NO:68)  
QKKDKTRWAEFREW-----NN (SEQ ID NO:69)  
GSRFSSPVGSTSLRLES-SDY--NY (SEQ ID NO:70)  
ADPSIYYSLXIBY-----KY (SEQ ID NO:71)  
DSPCYMPTMPAPPIRDSEFGW--DD (SEQ ID NO:72)  
TSSFYWYCTTAPPY-----NV (SEQ ID NO:73)  
TEIEWYGCNLRITTF-----TR (SEQ ID NO:74)  
NQLAGGWYLDPNYWLVSVGAY--AI (SEQ ID NO:75)  
RLTEMGACDARWATLATRTFAYNY (SEQ ID NO:76)  
DGWTRKKBGGIGLPWSVQCBDDGYNY (SEQ ID NO:77)  
DSVPCHL-----DV (SEQ ID NO:78)  
VBYPIDMCS-----RY (SEQ ID NO:79)

and/or,

in that its constant region comprises  $C_{H2}$  and  $C_{H3}$  domains comprising an amino acid sequence selected from the following sequences:

for the  $C_H2$  domain:

---

APETLGPTVFTEPPKPKDVLSTLTP (SEQ ID NO:31)  
 APELPGGPSVFEVPTKPKDVLSTSGRP (SEQ ID NO:32)  
 APELPGGPSVFEVPTKPKDVLSTSGRP (SEQ ID NO:33)  
 APELPGGPSVFEVPTKPKDVLSTSGRP (SEQ ID NO:34)

---

for the  $C_H3$  domain:

---

GQPREQVYTLA (SEQ ID NO:35)  
 GQPREQVYTLAPKRLLE (SEQ ID NO:36)  
 GQPREQVYTLPPSRDEI (SEQ ID NO:39)  
 GQPREQVYTLPPSRDEI (SEQ ID NO:40)  
 GQPREQVYTLPPSRDEI (SEQ ID NO:41)

---

and/or,

in that its hinge region comprises from 0 to 50 amino acids.

3. The immunoglobulin of claim 2, wherein said hinge region comprises a polypeptide having an amino acid sequence selected from the following sequences:

---

GTNEVCKCKPK (SEQ ID NO:37)  
 and,  
 EPKIPQPKPKPKPKPKPKPKPKPEPCKCKPK (SEQ ID NO:38)

---

4. An immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable ( $V_H$ ) region and a constant region, said constant region being devoid of first constant domain  $C_H1$ , wherein the immunoglobulin is devoid of polypeptide light chains, and wherein said immunoglobulin is of class 2 (IgG2) or class 3 (IgG3) and is obtained by a process comprising cloning DNA or cDNA sequences encoding an immunoglobulin or a  $V_H$  domain having a determined specific antigen binding site into the pMM984 plasmid and transfecting cells with the recombinant plasmid.

5. An immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable ( $V_H$ ) region and a constant region, said constant region

being devoid of first constant domain  $C_H1$ , wherein the immunoglobulin is devoid of polypeptide light chains, and wherein said immunoglobulin is of class 2 (IgG2) or class 3 (IgG3) and is obtained by a process comprising cloning DNA or cDNA sequences encoding an immunoglobulin or a  $V_H$  domain having a determined specific antigen binding site into a vector wherein the vector is a vector appropriate for expression in plant cells, and the transformed recombinant cells are plant cells.

6. The immunoglobulin of claim 5, wherein said vector is pMon530.

7. An immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable ( $V_H$ ) region and a constant region, said constant region being devoid of first constant domain  $C_H1$ , wherein the immunoglobulin is devoid of polypeptide light chains, and wherein said immunoglobulin is directed against a biological organism selected from the group consisting of a bacteria, a virus and a parasite.

8. An immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable ( $V_H$ ) region and a constant region, said constant region being devoid of first constant domain  $C_H1$ , wherein the immunoglobulin is devoid of polypeptide light chains, and wherein said immunoglobulin is directed against a biological molecule.

9. The immunoglobulin of claim 8 wherein said biological molecule is selected from the group consisting of a protein, a haptan, a carbohydrate, a nucleic acid, a cellular receptor, and a membrane protein.

10. An immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable ( $V_H$ ) region and a constant region, said constant region being devoid of first constant domain  $C_H1$ , wherein the immunoglobulin is devoid of polypeptide light chains, and wherein said immunoglobulin is conjugated with a toxin.

\* \* \* \* \*